Limit Order Book Resilience and Cross Impact Limit Order Book Model

by

Xin Geng

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Mathematics

February 2015

THE UNIVERSITY OF WARWICK

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this thesis are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other University. This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration, except where indicated in the text.

Acknowledgements

First and foremost, my sincerest appreciation goes to my supervisors, Dr. Florian Theil and Dr. Alex Stremme, for their thoughtful guidance, valuable support and encouragement, and patience in sustaining my motivation during the research and preparation of this thesis. Florian teaches me always to think in mathematically rigourous way, while Alex opens my mind on practical finance problems. Maybe most importantly, they strike a balance between mathematics and finance to give me room to develop my own ideas.

I thank Dr. John O'Hara and Dr. Oleg Kozlovski for accepting being my co-examiners. I greatly appreciate their comments.

I am also grateful to the staff members from Mathematics Institute in University of Warwick, especially Professor Dwight Barkley and Dr. Dmitriy Rumynin, for their huge support and smart suggestions. I would also like to thank my guardian angels during my study at Warwick, especially Lam, Chin, Adam, Hong, Tom, Jane, Simon and Mark. Their company, drinking party, lunch time chat, tuition on swimming all make me a better me.

Last but not least, I wish to express my deepest gratitude to my family. Without their infinite support and love, I will never accomplish this journey. Very special thanks go to Hengxu Wang, for his interesting stories when I soak in limit order books and for his support of my academic endeavors.

Abstract

This thesis comprises of five chapters. The first chapter gives a brief introduction on the existing literature about the optimal trading order execution problem, the concept of limit order book, market impact models and their underlying market microstructure. We will also provide some brief review on the regularity problem of market impact model and the resilience effect of the LOB market. Some notions about the limit order book trading will also be introduced in this chapter. The second chapter, a game theoretical model given by Rosu [74] is introduced and the same side and opposite side resilience are reinterpreted for this model. The solution structure of a Markov equilibrium of this model is obtained for the same side resilience by providing a rigourous mathematical analysis. We also provide a sufficient condition for the existence of real-valued solutions under this situation. We also reproduce the results in Rosu [74] about the opposite side resilience in this LOB model. In the third chapter, we extend the LOB market impact model in Obizhaeva and Wang [65] by introducing two sides resilience and a general LOB shape function. Two existing LOB market impact models are then replicated by our extended model, allowing the cross-impact resilience rate going to zero and infinity respectively. In the last two chapters, we conduct two applications of our extended market impact model. These two applications are able to help us study the optimal execution problem and the market regularity issues. We find out that the minimum cost of the zero-spread LOB model is a lower bound of the minimum cost of our extended market impact LOB model and those models with zero bid-ask spread have weaker regularity conditions than those with a non-zero bid-ask spread.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In section 1.1, we will briefly review the existing literature about the optimal trading order execution problem, the concept of limit order book (LOB), market impact models and the market microstructure underlying these models. Moreover, we will also provide some brief review on the regularity problem of market impact model and the resilience effect of the LOB market. Since our research is conducted on the limit order book structure, some notions about the order book trading will be introduced in section 1.2. These notions are used throughout this thesis. Then in section 1.3, we would like to present the motivation of this research. Finally, in section **??**, we will give the structure of this thesis and our main contributions.

1.1 Literature review

1.1.1 Optimal trading order execution problem

With or without private information, there are some cases in which a trader needs to liquidate a large amount of some asset. As proposed in Brunnermeier and Pedersen [19] some examples are hedge funds with margin calls, traders who uses portfolio insurance, stop loss options, or other risk management strategies, a short seller who may need to respond to the price increases[1]. Because of the limited liquidity in the financial market, trading a large amount of any asset has an impact on its prices, and usually this price shift is against the trader's interest. This trading incurred price shift is called the *price impact*.

The basic observation is that the costs of the price impact of a large trade can be reduced significantly by splitting the large trade into a sequence of smaller trades. To this end, we adopt the view of Bertsimas and Lo [13] and Gatheral [33] about the order execution process. They suggested that the trading process is separated into three layers: The first layer is called *macro-trader* and this layer decides about the timing of trading and about the order sizes; The second layer called *micro-trader*. Given a slice of the order placed by the macro-trader, the second layer decides whether to place market orders or limit orders. If a limit order is placed, at which price to trade; The third layer is the *smart order router*. Here a decision is made about which trading venue to send the orders to.

Although it is desirable to have an integrable model for all layers of the trading process, such a model might be overly complex as suggested in Bertsimas and Lo [13]. The *market impact models* we consider in this thesis are models developed for solving the first layer optimal execution problem. We will review the market impact models in detail in the section 1.1.2. Apart from the market impact models, there are other first layer related studies focusing on optimal trading times, such as Kharroubi and Pham [51] and Lehalle et al. [59], or the studies focusing on order split between transparent and hidden, such as Buti and Rindi [20], Cebiroğlu and Horst [22], Esser and Mönch [29] and Kratz and Schöneborn [55]. The second layer execution usually involves limit order placement, such as Avellaneda and Stoikov [10], Bayraktar and Ludkovski [11], Guéant and Lehalle [41], Guéant et al. [42], Harris [45], Hollifield et al. [47], Kovaleva and Iori [54]. There is also some research which combines both the first and the second layer, for example Guibaud and Pham

[1] As stated in Madhavan [62], this is the case that 'issues of how investors trade are decoupled from issues of why they trade'. Even though an ideal framework should combine the optimal trading strategy problem and dynamic portfolio problem, it is very helpful to have a full understanding on the trading part itself at first.

[43], Guilbaud and Pham [44], Huitema [50], Naujokat and Westray [64]. For the layer of smart order router, one can refer to Cont and Kukanov [24] and Laruelle et al. [57] and the references therein.

1.1.2 The market microstructure underlying market impact models

The market microstructure endogenously explain the formation of the price impact, and the motivation of trading activities, etc. To build a both mathematically tractable and well rounded market impact model, it is desirable to gain a better understanding of the endogenous market microstructure that underlies the market impact models. Some recent books and articles, such as Biais et al. [15], Lehalle [58], Madhavan [62], O'Hara [66], summarised the market microstructure literature from the aspects of theoretical, empirical and experimental study.

In a market impact model, the relation between the transacted order volume and the consequent price shift is described by the *price impact function*. The price impact function is an abstract microstructure description of the interactions between traders and the liquidity in the market. In other words, it does not model the dynamics of price impact via interactions of trades at a microscopic level as in a market microstructure model, but emphasise on a direct relation between a large order and the price dynamics. As in Alfonsi et al. [5], Almgren and Chriss [9], Gatheral [32] and Obizhaeva and Wang [65], the asset price in a market impact model is in the form of

$$S_t = S_t^0 + \text{Price impact function}$$

where S^0 is an exogenously given process to describe the asset price when it is unaffected by the price impact.

Both the unaffected price process and the price impact function are required to be specified in a market impact model. As suggested in Gatheral and Schied [35] and Schied and Slynko [78], there are two generations of market impact models. The *first generation market impact* models distinguish between two price impact components, namely the *temporary impact* and *permanent impact*. The *second generation market impact* models is based on the subsequently decay of price impact, which is called the *transient impact*. In the rest of this section, we will briefly review the market microstructure that underline the first and the second generations of market impact models.

The first generation market impact models and dealer market

Madhavan [62] attributes the cause of temporary and permanent price impact to three types of costs when one trades in a dealer market. On a dealer market, liquidity is provided by a specialist who is contractually obliged to always stand ready to buy at quoted bid and sell at quoted ask. Price is determined via the specialist's auction and the trader's bidding.

The temporary price impact reflects the transitory cost of demanding liquidity, such as order handling fee in Roll [72] and inventory cost in Stoll [81]. The temporary price impact only affects the individual transaction that has triggered it.

The permanent price impact reflects the specialist's price update based the information transmitted to the market by the buy/sell order flows. Thus the permanent price impact is due to the costs of being adversely selected by informed traders as discussed in Easley and O'Hara [28], Glosten and Milgrom [38], Kyle [56]. The permanent impact does not only influence the price of the current trade but also the prices of all subsequent trades.

This kind of market impact models are first introduced in Bertsimas and Lo [13], Almgren and Chriss [9] and Almgren [8]. The framework of Almgren and Chriss [9] has now been a basis for practical applications used in the financial industry. Some variants of the Almgren-Chriss framework are:

1. Adapting the optimal execution strategy for various risk criteria, such as the mean-variance optimisation utilised in Almgren [8] and Huberman and Stanzl

[49] and the expected utility maximisation applied in Schied et al. [79] and Schied and Schöneborn [77].

2. Introducing the optimal adaptive strategy, for example, Lorenz and Almgren [61];

3. Applying more general unaffected price processes, other than the Bachelier model, such as the geometric Brownian motion applied in the work of Schied [76] and Gatheral and Schied [34].

The second generation market impact models and LOB market

On a LOB market, there is no designated liquidity provider. Liquidity is offered in a self-organised way. In other words, any agent can choose, at any instant of time, to either provide liquidity or consume liquidity. A satisfactory LOB microstructure model should be able to include the interactions of different traders and explain how the bid and ask prices are affected by the interactions. Particularly in this thesis, it is desirable to reflect the same side and opposite side resilience effect in a LOB model.

We list some survey papers focusing on the market microstructure of the LOB. They are Bouchaud et al. [18], Gould et al. [40] and Parlour and Seppi [68]. Following the view of Abergel and Jedidi [1] on the classification of LOB microstructure models, two research methods could be outlined. The first method is a game theory approach, such as Foucault et al. [30], Goettler and Rajan [39], Parlour [67], Rosu [74] and Rosu [75]. Among these, the effect of asymmetric information on LOB trading activities is considered in Bloomfield et al. [16], Goettler and Rajan [39], Harris [45] and Rosu [75]. In such a trading game, traders are assumed to arrive at the LOB market randomly and trade strategically by endogenously choosing their trading decisions as solutions to individual utility maximisation problems. The LOB dynamics is then the collections of equilibrium strategies of all active traders.

The other method focuses on modelling the LOB microstructure in the zero-intelligence models, such as Abergel and Jedidi [1], Cont and Larrard [25], Cont and Larrard [25], Smith et al. [80] and Toke [82]. Zero-intelligence means the focus of this approach is more on reproducing the mechanics properties of the order book without assuming the strategic interactions between agents. The arrivals of different order flows are assumed to be independently and identically distributed point processes.

More empirical studies suggest that price impact is transient, but is not separated into two parts, namely the permanent price impact and temporary price impact. In other words, an order creates some immediate price impact that subsequently decays over time. Bouchaud et al. [17], Bouchaud et al. [18], Potters and Bouchaud [70], Weber and Rosenow [85], Wyart et al. [87] are among those empirical LOB studies which support this transient decay idea. Obizhaeva and Wang [65] adopted the transient price impact with the same side resilience into the market impact models. In their research, the LOB is defined via a block shaped and time independent function. The same side resilience factor was defined via a deterministic exponential function. In recent years, there are some research extended the work of Obizhaeva and Wang [65]:

1. In Gatheral et al. [37] and Alfonsi et al. [6], the same side resilience factor follows other deterministic functions.

2. In Alfonsi et al. [4], Alfonsi and Schied [3], Predoiu et al. [71] and Alfonsi et al. [5], the shape function of the LOB is also defined as a time independent function but not in block shape.

3. In Alfonsi and Acevdeo [2] and Fruth et al. [31], the LOB shape function is assumed to be a block shaped function but not independent with time.

4. In Fruth et al. [31] and Weiss [86], the same side resilience factor is considered to be stochastic.

5. More general unaffected price process is applied in Lorenz and Schied [60].

1.1.3 Regularity of market impact models

The optimisation problem brings us the issue of regularity of a market impact model. A minimal regularity condition is the existence of admissible optimal execution strategy. This existence of an optimal solution is guaranteed by the absence of the *price manipulation strategy* (PMS) and the *positivity of liquidation cost* (PLC). Moreover, the resulting optimal strategies should be well-behaved. For instance, we do not want to follow a trading strategy that strongly oscillates between buy and sell since there is usually additional fees for trading market order. This oscillation strategy can be excluded by the absence of the *transaction-triggered price manipulation* (TTPM) in a market impact model. The notions of these irregularity conditions of market impact model will be introduced in section 3.3.

However, we should note that these arbitrage opportunities in market impact models are different from the arbitrage in derivative pricing models and are also distinguished from the arbitrage opportunities generated by asymmetric information as discussed in Allen and Gorton [7]. So even a martingale assumption on the unaffected price process will not exclude these irregularities in market impact models.

Huberman and Stanzl [48] initiate the research on arbitrage opportunity in market impact models and link this to the arbitrage in derivative pricing framework via so called quasi-arbitrage. They find that the linearity of the permanent price impact function is a necessary condition for the absence of the PMS. Their results are further confirmed in transient models of Gatheral [32] and Gatheral et al. [36]. Alfonsi et al. [6] and Gatheral et al. [37] both research on linear transient LOB model with general decay factor. They find out a necessary condition on the decay factor for absence of both PMS and TTPM. Alfonsi and Schied [3] and Klöck [52] represent two extensions of the linear transient impact model. Alfonsi and Schied [3] generalise their results with the non-linear price impact function. Klöck [52] introduce the effect of stochastic linear price impact into market irregularity investigation and propose the notion of PLC.

1.1.4 The resilience of the market impact

In this section, we provide an overview of some theoretical and empirical studies on the order book resilience. In market impact modelling framework, the resilience means the price impact induced by a large trade can be reduced if extra time is given for the market to recover. The resiliency in a LOB is quantified in three respects:

1. Magnitude, i.e. how much will the best bid or ask price recover.
2. Speed, i.e. how quick will the best or ask price recover.
3. Direction, i.e. same side resilience or opposite side resilience[2].

In particular, we will have a discussion on the same side and opposite side resilience after a trade incurs price impact.

Theoretical studies on the resilience of the market impact

The resilience study in Bouchaud et al. [17] and Bouchaud et al. [18] is conducted via analysing the diffusive behaviour of mid-price (average of the best bid price and the best ask price) process. They claim that the market observed mean-reverting mid-price process holds if and only if the price impact decays slowly. They offer the rational behind price mean-reverting as: the liquidity provider needs to close their position at a later time without trading at a much higher price. So the liquidity provider needs to mean-revert their limit order prices after a large order. The recovery direction is not clarified and their view about the time scale of resilience is that the mean reversion cannot take place too quickly.

Via a study on high frequency (HF) trader's optimal market making (i.e. supplying liquidity with limit orders) strategy, Cartea et al. [21] offer evidence on opposite resilience in a LOB market on short time scale. The optimal solution implies that the HF trader should correct the market making strategy in order to

[2]See section 1.2 for a detailed discussion

mitigate the risk of being adverse selected, and exploit the short-term mid-price deviation via directional strategy. In other words, after a large market buy order, the HF trader posts further away from the mid-price on the ask side (avoid being adversely selected) and puts closer to the mid-price on the bid side (exploiting the short-term deviation by directional strategy). That is to say, HF market maker's directional strategy supports the opposite side resiliency of the order book on short time scale.

Foucault et al. [30] applied a game theory to study continuous-time order submission problem. To measure the resiliency they use two definitions, namely the reservation spread and the competitive spread. Reservation spread is the smallest price improvement by which a limit order trader could make a non-negative profit. Competitive spread is the patient traders' reservation spread. Resiliency is then measured as the probability that the spread reverts to 'competitive spread' before the next transaction. They find out that when traders have the same reservation spread, the resilience is the fastest and when traders are heterogeneous, the more the patient traders there are, the faster the price impact decays. There is no clear results about the resiliency direction.

Empirical studies on resilience of the market impact

Ponzi et al. [69] and Tóth et al. [83] are two empirical studies via event study based on an order book. They both observe a power lay decay of price impact. In other words, new limit orders are not placed simply in a way that immediately reverts the spread back to its original value, but they are sequentially placed close to the current best price and this leads to a slow decay of the spread. However, the study of Ponzi et al. [69] does not distinguish the resilience between the same side and the opposite side recovery. Tóth et al. [83] separately measure the number of queuing orders on each side of the order book before and after a large market order. They show for price drop on the bid side the number of queuing limit buy orders decreases to about half of the usual value. At the same time on the ask side the number of queuing limit sell orders increases and even stays very high for a long time after the event. That is just the two-side resilience effect which will be explained in more detail in next section.

Another two important empirical study on order book resilience are: Biais et al. [14] and Degryse et al. [27]. Biais et al. [14] find new in-spread limit orders on the ask (bid) side of the market are particularly frequent after large market sell (market buy) orders. This empirical observation is an evidence of opposite side resilience. They attribute the co-movement effect to information, in the sense that part of the resilience in the same side could be mechanical, but the resilience in the opposite side is believed to be due to the information arising from the shift in the expected fundamental value. Degryse et al. [27] extend the work of Biais et al. [14] by studying not only the next incoming limit order after the large market order but also the sequential orders. They conduct a descriptive event study and they observe that for small stocks the best ask price jumps up after a large market buy order, while the best bid price increases as well but without a jump. So does after a large market sell order. The resilience of two sides of the order book are different to the same large market order..

1.2 LOB and its execution rules

In most of the published literature on LOB, the following terminology has been adopted. They will be widely used in this thesis too.

Market order and limit order

A limit order is an order to buy or sell some specified quantity of an asset at a price specified by the trader. It is not executed immediately, instead it enters the queue of outstanding limit orders. A market order of a given size is an order that results in an immediate execution at the best available price upon submission. Thus for market order traders, they only specify the amount to buy or sell without explicitly specifying a trading price. Market order is executed against the existing limit orders. This is the interaction between market and limit orders.

Large order and child order

According to an empirical paper Biais et al. [14], trades are distinguished by trading direction (buy or sell) and aggressiveness (patience for order execution). The most aggressive order is placed via a market order to trade a large quantity of shares which usually demand the execution of more limit orders than that available at the best price.

In market impact models, a large order corresponds the most aggressive order in market. In particular, it specifies a large amount of shares which is needed to be split into smaller orders, so that the adverse price impact incurred by it can be reduced. This sequential smaller order is called the child order.

Bid side and ask side

A LOB for a single asset is the collection of all active buy and sell limit orders with the corresponding prices and volumes information. We call the collection of limit buy orders the *bid side* of the LOB and the collection of limit sell orders the *ask side* of the LOB. A snapshot of an order book is shown in Figure 1.1.

Order execution rules

Execution of outstanding limit orders by market orders is settled according to a set of priority rules. Prevailing ones are price-time rule and pro-rata rule.

In this thesis, it is assumed that the LOB follows the price-time priority rule. Price-time priority LOB markets give priority to orders submitted at more competitive prices and to displayed orders over hidden orders at the same price level. Orders with the same display status and submission price are usually matched on a

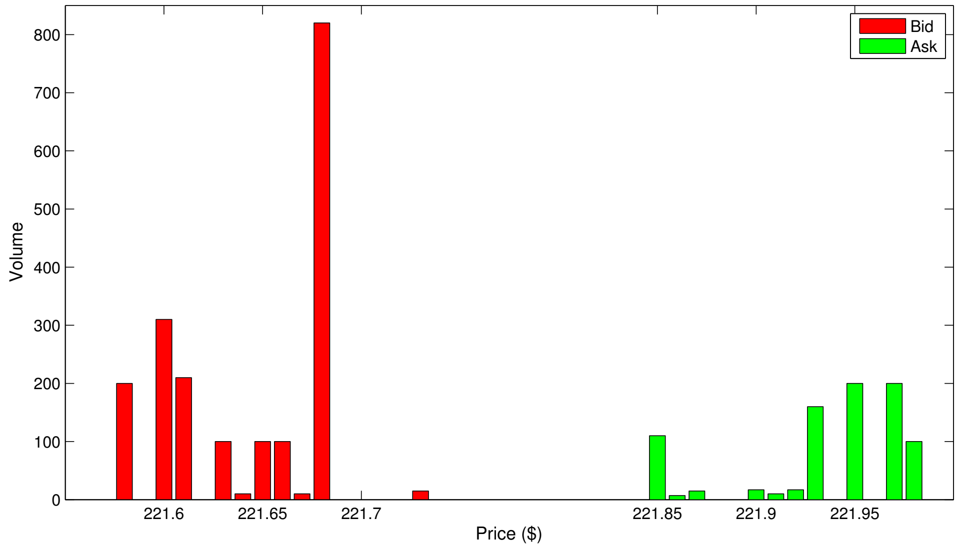


Figure 1.1: A LOB Snapshot of AMZN on June 21, 2012.

first-come-first-serve basis. As a result, the price-time priority order execution rules are:

1. When a market buy order arrives, limit sell orders in the ask side of the LOB are executed, starting from the orders with the lowest price to the more expensive ones until the total number of shares ordered is reached.

2. When a market sell order arrives, limit buy orders in the bid side of the LOB are executed, starting from the most expensive orders to the less expensive ones until the total number of shares ordered is reached.

Discrete price grids

In a real LOB, orders are placed at a pre-fixed discrete price grids. The grid step is the smallest interval between two prices and it is called the *tick size*. The tick sizes are different between the exchanges and trading assets.

LOB dynamics and order flow

The order book evolves over time according to the arrival of new orders. The price is therefore the result of the interactions between the order book and order flow. The dynamics of the LOB would be mainly affected by the following four order flow processes, namely market orders, limit orders placed in the bid-ask spread, limit orders placed at prices worse than the best price and cancellation of limit orders.

Same side resilience and opposite side resilience

The same side resilience describes the following market situations:

1. After a market buy order, the ask side of the order book will be recovering to its original status.
2. After a market sell order, the bid side of the order book will be recovering to its original status.

Likewise, the opposite side resilience describes the following market situations:

1. After a market buy order, the bid side of the order book will be recovering to its original status.
2. After a market sell order, the ask side of the order book will be recovering to its original status.

In terms of order flow, the same side resilience means that:

1. After the arrival of a market buy order, there are new incoming limit sell orders placed within the current bid-ask spread.
2. After the arrival of a market sell order, there are new incoming limit buy orders placed within the current bid-ask spread.

Likewise, in terms of order flow, the opposite side resilience means that:

1. After the arrival of a market buy order, there are new incoming limit buy orders placed within the current bid-ask spread.
2. After the arrival of a market sell order, there are new incoming limit sell orders placed within the current bid-ask spread.

Two sides resilience means that, after the arrival of any market order, there are new incoming limit buy and sell orders placed within the current bid-ask spread.

1.3 Motivations

As demonstrated in the section 1.1.4, empirical evidences of the same side and opposite side resilience of the market impact are discovered in trading actives. We believe that these resilience effects are important academically and practically. But because of the complexity of these features, they are still not well researched. In Rosu [74], a microstructure LOB model is provided and to the best of our knowledge, among other LOB models, this model can theoretically replicate the effect of two-side market impact resilience due to its mathematical tractability.

Different from the microstructure LOB model provided by Rosu [74], a market impact model based on the LOB is introduced by Obizhaeva and Wang [65]. As reviewed in the last part of the section 1.1.2, the work of Obizhaeva and Wang [65] is very popular in recent years and there are various modifications or generalisations conducted by other academics. However, all these studies did not consider two sides resilience. Usually, they simplify their models by zero-spread assumption or focusing on only one-side of the order book.

This thesis fills a gap in literature by providing a mathematically rigourous proof for the existence of the same side and the opposite side resilience in the microstructure LOB model of Rosu [74]. Moreover, to the best of our knowledge, there is no published evidence which examines the effects of the two sides resilience in a market impact model. As a consequence, this thesis also fills another gap that considering two sides resilience effects with a non-zero bid-ask spread under the Obizhaeva and Wang [65] market impact model.

After providing a generalised Obizhaeva and Wang [65] market impact model, which we call a cross-impact LOB model, research on the optimal execution problem and the market regularity issues under this model is conducted for two reasons. Firstly, Klöck [52] and Fruth et al. [31] claimed, but did not prove, that the market impact models with zero bid-ask spread has weaker regularity conditions than the models with a non-zero bid-ask spread. So, we would like to verify this argument in our cross-impact LOB model. Secondly, we want to investigate the effects of two sides resilience on the first layer optimal execution strategy, which has long been neglected in the literature.

Chapter 2

Resilience in LOB microstructure model

In this chapter, our aim is to show the presence of same side resilience and opposite side resilience in an order book market. We base our analysis on a game theoretical model given by Rosu [74]. Compared to other game theoretical models summarised in Section 1.1.2, there are several reasons to work on the model of Rosu [74]. First, it allows for a flexible spread of the order book and for explicit measurement of resiliency with enough mathematical tractability, compared to the models of Parlour [67], Foucault et al. [30] and Goettler and Rajan [39]. Second, the direction of the resilience is able to be reflected and modelled in this model, while it is assumed a fixed one tick size spread in Parlour [67]. Third, the motivation of trading is exogenous to the model, namely the effect of asymmetric information is not modelled. Rosu [75] provides an alternative model in which the information effect is considered. Fourth, it includes the execution time risk of limit orders into the model formulation, as in Foucault et al. [30].

In Section 2.1 we summarise the model of Rosu [74]. It starts by reviewing the main characteristics of this continuous time trading game proposed in Rosu [74]. We then introduce the construction of the rigid, competitive Markovian equilibrium in the one-side book case, since this case is quite intuitive and admits closed-from equilibrium strategies. It turns out that the properties of one-side case can be easily extended to the two-side case.

We present our main results in Section 2.2. In Section 2.2.1 we reinterpret the same side resilience and opposite side resilience by notions of Rosu [74]'s trading model. The same side resilience is reflected by the positivity of the difference between same side temporary price impact and permanent price impact caused by a same side market order. The opposite side resilience is measured by the positivity of the non-execution side permanent price impact. This is an essential step since the price adjustment in this continuous time trading game is taking place instantaneously, while the transient resilience is a time-dependent feature of a LOB. Section 2.2.2 deals with the solutions of the Markov equilibrium obtained in the one-side order book case. We present the solution structure in our Proposition 2.2.2 which is distinguished from the results of Rosu [74]. In Section 2.2.3, we provide a rigourous proof of the same side resilience by looking at the asymptotic behaviour of the price impact functions under the assumption of fast decay arrival rates. We also construct a counterexample of the same side resilient in Proposition 2.2.4 where the fast decaying assumption did not hold. Section 2.2.4 reproduce the results in Rosu [74] about the opposite side resilience in this LOB model.

2.1 The model description

2.1.1 Characteristics of the LOB trading game

The limit order book is of one asset with no dividend. The game is happening on time interval $[0, \infty)$, and trading takes place in continuous time[1]. The tick size of price grid is zero, i.e. the prices can take any real value. We take the reason to trade for all traders to be exogenous to the model. All information about the order

[1]Since there is no universally accepted model of continuous time stochastic game and related definitions, we follow Rosu [74] that adopts and extends the continuous time framework of Bergin and Macleod [12] and Rosu [73].

book and trader strategy are publicly available. Traders can choose between market order and limit order. The execution of limit orders is subject to the price-time priority rule as discussed in the Section 1.2. The limit order can be cancelled or revised at will with no cost. After order execution, the agents exit the order book forever. There is no delay in trading, in the sense that submitting or execution happens instantaneously.

The players form a countable set. They are distinguished by the patience and trading directions of buy or sell. More precisely, there are four types of traders considered here: patient buyer, patient seller, impatient buyer, and impatient seller. Each player's choice of buy or sell, patience and amount of trading orders are exogenously given outside of the model and stay the same during the trading game. For simplicity, it is assumed that the impatient traders only submit market order, which is automatically executed against existing limit orders. Patient traders choose strategically between market and limit order and the price to submit if a limit order is used. The price of a limit sell order, P^S, and the price of a limit buy order, P^B, is constrained to lie between $[\underline{P}, \overline{P}]$. Moreover, it is assumed that there is an infinite supply (demand) at price $\overline{P}$ ($\underline{P}$) which is provided exogenously. Patient traders arrive with only one unit to trade. Impatient traders can submit up to k-unit market orders for some constant $k \geq 1$.

The arrivals of traders at the order book are modelled according to independent Poisson processes with constant, exogenous intensity rates. Denote the arrival rates for different type of traders as following:

$$\begin{cases} \text{For } i \leq k,\ \mu_i > 0 \text{ is the arrival rate of } i\text{-unit impatient buyer. If } i > k,\ \mu_i = 0; \\ \text{For } i \leq k,\ \lambda_i > 0 \text{ is the arrival rate of } i\text{-unit impatient seller. If } i > k,\ \lambda_i = 0; \\ \mu > 0, \quad \text{arrival rate of patient seller;} \\ \lambda > 0, \quad \text{arrival rate of patient buyer.} \end{cases}$$

The utility function is a trade-off between higher execution price and lower waiting costs. More specifically, the utility function for market order traders is the instantaneous best prices at the submitting time, since they do not need to wait for execution. The expected utility functions for limit order seller and buyer at time $t \in [0, \infty)$ are defined respectively as

$$f(P^S, t) := \mathbb{E}_t[P^S - r(\tau - t)]$$

and

$$g(P^B, t) := -\mathbb{E}_t[-P^B - r(\tau - t)],$$

where $\tau > t$ is the random execution time and r is the waiting cost discount factor for patient agents. That is, limit order traders bear the execution time risk. There are a few things we should note about these utility functions. First, $-g$ represents the expected utility of a buyer since the higher utility is obtained for a buyer by lowering the price to pay. The sign '$-$' was moved to the right hand side so that it is convenient to maximise both f and g for buyer and seller at any time. Second, by assuming a linear relationship between the execution price and the discounted waiting cost, there is possibility that the utility becomes negative. One could instead consider an alternative form of utility functions

$$f(P^S, t) := \mathbb{E}\left[{}^S P e^{-r(\tau - t)}\right]$$

and

$$f(P^B, t) := \mathbb{E}\left[P^B e^{-r(\tau - t)}\right].$$

However, the model is already too complicated to add more complexity from the utility function. To this end, we can always choose the waiting cost discount factor r such that the utility is positive. In fact, we will see in the sequel that in equilibrium the utility has a positive constant lower bound.

The players can respond immediately in this continuous time trading setting. In other words, the trader will keep the current order as long as the other agents stays unchanged, but if some other agents deviate from the current state at some time t, this trader will immediately undercut her order price at t.

2.1.2 One-side order book game

In this section, we will consider firstly one-side of the order book. Without loss of generality, we focus on the ask side of a LOB. For the ask side, the best bid price is the reservation value of the limit order sellers, as they can always submit a market sell order at the best bid price and exit the trading. So we set the lowest price $\underline{P}$ to be the best bid price for sellers.

For description simplicity, we set $\Sigma := \lim_{k\to\infty} \sum_{i=1}^{k} \mu_i$ and introduce the following notations:

m: the number of limit order sellers in the order book;

$a_m(i)$: the ask price of the ith limit sell order counted from the best ask price, in an order book with m limit order sellers with $i = 1, ..., m$;

a_m: the best ask price for a m limit order seller order book.

Firstly, we review the construction of the competitive Markovian equilibrium in Rosu [74]. At the beginning, the order book is empty, i.e. $m = 0$. At some time t patient seller 1 comes and places a limit order at the maximum price level $a_1 = \overline{P}$. This is required by the Nash equilibrium since $\overline{P}$ is the highest price as long as seller 1 is the only one in the order book. Suppose then a second patient seller 2 arrives. If immediate response is not allowed, seller 2 would secure earlier execution by placing a limit order at $a_2 = a_1 - \delta$ for some very small $\delta > 0$. The limit order of seller 2 gets executed before that of seller 1 with only infinitesimal price sacrifice, the expected utility of seller 2 is strictly higher than that of seller 1. However, seller 1 is allowed to undercut his limit order instantaneously to $a_1 - 2\delta$. So a price war would follow. As the limit order price gets undercut, the utility for both of them decreases. Therefore, in equilibrium, seller 2 should submit a limit order at $a_2 < a_1 = \overline{P}$ such that both sellers have the same expected utility.

The same idea works for m-seller order book as well. In equilibrium, the sellers have their limit orders placed at different prices but get the same expected utility; otherwise, they would be undercut by each other. The same-utility property reflects the assumption that the seller with a higher limit order price needs to wait longer. As a result, we could show that in equilibrium the utility of limit order seller depends only on the number of active sellers in the book m. Thereafter, we denote the seller's expected utility by f_m. Since the arrival of patient sellers is modelled by exogenously given independent Poisson processes, the state variable m is exogenous and follows a Markov process. According to the definition of Markov strategy in Maskin and Tirole [63], the state m utility function f_m and the trading strategy form a Markovian perfect equilibrium.

Apart from determining the price strategies in equilibrium, we also need to know how limit sell order traders revise their orders when someone in the book deviates. To this end, Rosu [74] applied the notion of competitive[2] equilibrium, in which a local deviation from one trader can be stopped by any other trader's immediate undercutting, assuming that the rest of the equilibrium does not change. The features of large amount trader and instantaneous response make the choice of competitive equilibrium natural[3].

Then we consider some properties of the utility function f_m, i.e. the boundary conditions of f_m. We should note that the state variable m is the only state member and must be finite. Suppose $m \to \infty$, the expected execution time for the top seller, i.e. the seller with the highest ask price, would be $\mathbb{E}[\tau] \to \infty$. Hence the top seller's expected utility is $\mathbb{E}[a_m(m) - r(\tau - t)] \to -\infty$. Instead, the top seller could at least improve her trading by submitting a market order with price $\underline{P}$ which gives him the reservation utility.

[2]An example of non-competitive equilibrium is given in the NASDAQ dealer market study by Christie and Schults [23].

[3]By Rosu [73], an Markovian equilibrium is competitive if a restriction of the strategies on time interval $(t, t+\delta)$ is still a Markov equilibrium. This restriction can be made because of the Markov condition with which pay-off related history is reduced to the limit of outcomes at a single time point.

In addition, from the economic point of view, we obtain that the expected utility function f_m decreases with m and the best ask price a_m is decreasing with m (since more limit order sellers incur longer waiting time). Meanwhile, the minimum value of a limit sell order should be no less than the reservation value $\underline{P}$. In other words, when the utility of waiting till execution is less than the utility of submitting a market sell order at $\underline{P}$(recall that $\underline{P}$ is defined to be the best bid price), some seller would just cancel their limit orders and place market orders at $\underline{P}$. Then it is reasonable to define a maximum capacity $M := \max\{m : f_m \geq \underline{P}\}$[4] and a *mixed strategy* in which limit order trader randomly switches to a market order with Poisson(ν). There are four types of mixed strategies that could happen when there are M sellers. These are introduced in Proposition 12 and Corollary 2 in Rosu [74]. In particular, the notion of rigid equilibrium is constructed here. Rigid equilibrium means that if some agents have mixed strategies, mixing is done only by the agents with the most competitive limit orders (highest bid or lowest ask).

Then, we derive a recursive system of utility functions f_m to compute the equilibrium strategy. For a book of $m < M$ sellers in equilibrium, the market can go to: state $m+1$ if a new limit order seller arrives with probability $\frac{\mu}{\mu+\Sigma}$, or to state $m-i$ with $i = 1, ..., k$ if an i-unit market order buyer arrives with probability $\frac{\mu_i}{\mu+\Sigma}$, where $\Sigma = \sum \mu_i$. Apart from this, patient sellers lost utility in a way proportional to expected waiting time with discount factor r. One obtains the formula

$$f_m = \frac{\mu}{\mu+\Sigma} f_{m+1} + \sum_{i=1}^{k} \frac{\mu_i}{\mu+\Sigma} f_{m-i} - \frac{r}{\mu+\Sigma}.$$

Furthermore, we construct the individual limit sell order prices $a_m(i)$ for $i = 1, \ldots, \min\{m, k\}$. With rate μ_i, an impatient buyer arrives and places an i unit market order. The ith seller will receives $a_m(i)$. The non-executed sellers will have

[4]If $f_M < \underline{P}$, some of the sellers could always improve their utility by submitting a market order directly at price $\underline{P}$ and get a better utility $\underline{P} > f_M$. This is a contradiction with the optimisation of each trading strategy. If $f_M > \underline{P}$, let us consider what happens in state $M + 1$. If one agent accepted $h = \underline{P}$ at some time and exited the game, the utility of the other agents would be $f_M > \underline{P}$. The utility for the agent who accepts h is lower than f_M. So no seller would accept h and everybody waits. But this is in contradiction with M being the largest state in which agents wait.

the same utility which is equal to f_{m-i} with probability $\frac{\mu_i}{\Sigma}$. Since all sellers must have the same expected utility in equilibrium, the ith price in a m-seller order book $a_m(i)$ must be equal to the expectation of all possible utilities, i.e.

$$a_m(i) = \frac{\sum_{j \geq i} \mu_j f_{m-j}}{\sum_{j \geq i} \mu_j}. \tag{2.1.1}$$

Define by convention that $f_i = \overline{P}$ for $i \leq 0$.[5]

When the order book reaches its maximum capacity M, the arrival of a new patient seller does not affect the state of the order book since in equilibrium the new arrival will immediately place a market order at price $\underline{P}$ and exit. With probability $\frac{\mu_i}{\nu+\Sigma}$ an i-unit market order buyer arrives, the system would go to state $M-i$. Or with probability $\frac{\nu}{\nu+\Sigma}$ the bottom seller, i.e. the seller with the lowest ask price, switches to a market order at $\underline{P}$ and exits by the rule of rigid equilibrium. Similarly one obtains a formula for state M utility, which is

$$f_M = \frac{\nu}{\nu+\Sigma} f_{M-1} + \sum_{i=1}^{k} \frac{\mu_i}{\nu+\Sigma} f_{M-i} - \frac{r}{\nu+\Sigma}.$$

Theorem 2 in Rosu [74] provides the strategy in equilibrium: if $m = 1$, then place a limit order at $a_1(1) = \overline{P}$; If $m = 2, ..., M-1$, look at the bottom k levels (or at all m levels if $m < k$), which are $a_m(1), ..., a_m(k)$. If any of them is not occupied, occupy it. Anything above $a_m(k)$ does not matter; if $m = M$, the strategy is the same as for $m = 2, ..., M-1$, except for the bottom seller at $a_M(1)$, who exits (by placing a market order at $\underline{P}$) after the first arrival in a Poisson process with intensity ν; If $m > M$, then immediately place a market order at $\underline{P}$. The equilibrium is unique in the class of rigid equilibria, in the sense that any other rigid equilibrium leads to the same evolution of the state variables.

[5] This can be seen by equation (2.1.1) and $a_1 = \overline{P}$. It is technically required that $a_1 = a_1(1) = \frac{\sum_{j \geq 1} \mu_j f_{m-j}}{\sum_{j \geq 1} \mu_j} = \overline{P}$, although for $i \leq 0$ f_i has no practical meaning.

2.1.3 Two-side order book game

Now we are ready to review the two-side case in Rosu [74], since the derivation of the opposite-side resilience depends on the knowledge of the best prices $a_{m,n}$ and $b_{m,n}$. Similarly as in the one-side case, we will familiarize ourselves with the possible trading activities, the set of state variables (m, n) where m is the number of patient sellers and n the number of patient buyers in the book, and the recursive system of utility functions $f_{m,n}$ and $g_{m,n}$. Thereafter, if given a solution of the recursive system of f, g, according to Definition 4 in Rosu [74] one can then obtain the formulas for best prices $a_{m,n}$ and $b_{m,n}$.

For simplicity in the two-side game, we assume that the impatient traders can only submit one-unit order, i.e. $\mu_i = 0$ and $\lambda_i = 0$ if $i > 1$, and will denote by

m **:** the number of limit order sellers in the order book;

n**:** the number of limit order buyers in the order book;

$a_{m,n}$**:** the best ask price in a m limit sell and n limit buy order book;

$b_{m,n}$**:** the best bid price in a m limit sell and n limit buy order book.

For illustration, let us see the trading activities for example at state $(1, 0)$. The market may go to:

$(0, 0)$ if either an impatient buyer arrives after time $T_1 \sim \exp(\mu_1)$ and buy with a market order at price $a_{1,0} = \overline{P}$, since $a_{1,0} = \overline{P}$ is the highest possible sell price to achieve for the patient seller in the range $[\underline{P}, \overline{P}]$; or after some random time $T_2 \sim \exp(\lambda)$ a patient buyer arrives and submits a limit buy order at price h. The existing patient seller accepts to trade at h if the expected utility of waiting till next arrival of a new agent is less than h according to the Proposition 12 in Rosu [74].

$(1, 1)$ if after some random time $T_2 \sim \exp(\lambda)$ a patient buyer arrives and submits a limit buy order at price h, but the existing seller does not accept h. Then the patient buyer stays and behaves as a monopolist at the bid side by changing bid price from h to a lowest buy price at B immediately.

$(2,0)$ if another impatient seller arrives after time $T_3 \sim \exp(\lambda_1)$.

As in the one-side case, if an order book with m limit order sellers and n limit order buyers is in equilibrium, all the m sellers must have the same expected utility denoted by $f_{m,n}$, and all the n buyers must have the same (minus) expected utility denoted by $g_{m,n}$. It is defined that the state region Ω as the set of all states (m,n) where in equilibrium agents wait in expectation for some positive time. The state variable m,n are finite because of the same reason as stated in the one-side case.

Thereafter, we can summarize all possible trading activities. For any state $(m,n) \in \Omega$, the system can go to the following neighbouring states:

Activity A**:** leading to $(m-1,n)$, (1) if an impatient buyer arrives, or (2) if a patient buyer arrives and place a market order to trade with the bottom limit order seller at the best ask price $a_{m,n}$, or (3) if a patient seller cancels the limit order and submits a market order at $\underline{P}$ when $n = 0$;

Activity B**:** leading to $(m+1,n)$, if a patient seller arrives and submits a limit order;

Activity C**:** leading to $(m,n-1)$, (1) if an impatient seller arrives, or (2) if a patient seller arrives and place a market order to trade with the top limit order buyer at the best bid price $b_{m,n}$, or (3) if a patient buyer cancels the limit order and places a market order at $\overline{P}$ when $m = 0$;

Activity D**:** leading to $(m,n+1)$, if a patient buyer arrives and submits a limit order;

Activity E**:** leading to $(m-1,n-1)$, (1) if the existing bottom patient seller switches to a limit order at a lower than best ask (bid) price h such that $h = f_{m,n} = g_{m,n}$, and the existing top patient buyer immediately accepts it by placing a market order; (2) if the existing top patient buyer switches to a limit order at a higher than best bid price h such that $h = f_{m,n} = g_{m,n}$, and the existing bottom patient seller immediately accepts it by placing a market order.

Now we are ready to define the mixed strategies under this two-side case. First of all, it is observed from Proposition 12 case (1) that it must be $f_{m,n} > g_{m,n}$ if all traders prefer to wait in the order book. Otherwise, in state (m, n) with $f_{m,n} < g_{m,n}$, the patient sellers could improve their utility by placing a limit order at some price level $h \in [f_{m,n}, g_{m,n}]$ and some patient buyer immediately accepts one of the offers by placing a market order. This is feasible by noting that $-g_{m,n}$ is the expected utility for buyers. This reflects the fact that as new patient buyers and sellers arrive, they place limit orders on both sides until it is better off to trade immediately rather than wait. In case of state (m, n) with $f_{m,n} = g_{m,n}$, the order book is defined to be full and the traders on both sides play a game of attrition. By Proposition 12 and Corollary 2 in Rosu [74], according to a Poisson process with intensity $\nu_{m,n}$ traders on both sides might apply mixed strategies. More specifically, they are one of the following four cases of mixed strategies: A(3), C(3) and E(1), E(2). The set of the states (m, n) at which mixed strategies take place is defined as the boundary γ of Ω.

In the two-side limit order book, some typical points in Ω can be summarised as in Figure 2.1, in which the big dots are boundary and small dots are the interior of the state region. Notations in this figure follow those in Rosu [74]. We should note that Ω does not have to be that specific shape.

For each point (m, n) in the state region, we can get a coupled system for $f_{m,n}$ and $g_{m,n}$ and corresponding equations for $a_{m,n}$ and $b_{m,n}$. Recall that k is the maximum units that a market order can trade with and r is the waiting discount factor. For each boundary point $(m, n) \in \gamma$, a corresponding number $\nu_{m,n} \geq 0$ is the intensity of a Poisson process with which some traders will have mixed strategies. Let ν be the collection of all $\nu_{m,n}$. Set $\Sigma_1 = \mu + \lambda + \mu_1 + \lambda_1$, and $\Sigma_2 = \mu + \lambda + \lambda_1$.

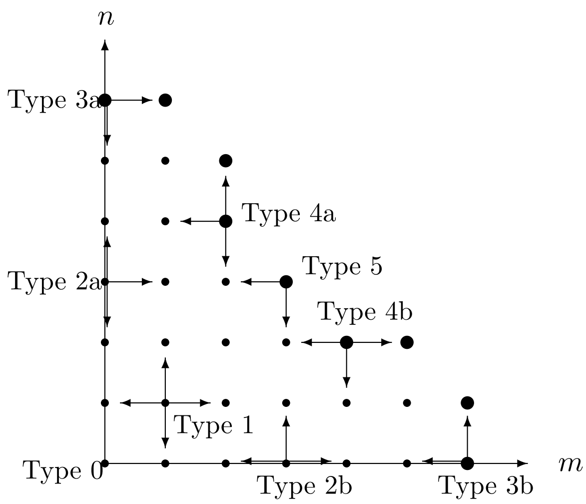


Figure 2.1: Types of points in the state region Ω cited from Rosu [74]. The points at boundary are types $3a, 3b, 4a, 4b, 5$. The types $0, 1, 2a, 2b$ are interior points.

Then define the recursive system associated to (Ω, k, ν) by considering the following set of equations for each state $(m, n) \in \Omega$.

First, we consider the interior points at which there are no mixed strategies. If (m, n) is of type 0, as illustrated in Figure 2.1, we have $m = 0$ and $n = 0$. Then we define by convention $f_{0,0} = \overline{P}$ and $g_{0,0} = \underline{P}$. If (m, n) is of type 1, there are four possible trading activities which can affect the state. They are activities $A(1)$, $C(1)$, B and D. One obtains the recursive equations

$$f_{m,n} = \frac{1}{\Sigma_1} \left[f_{m+1,n}\mu + f_{m,n+1}\lambda + f_{m-1,n}\lambda_1 + f_{m,n-1}\mu_1 - r \right]$$

and

$$g_{m,n} = \frac{1}{\Sigma_1} \left[g_{m+1,n}\mu + g_{m,n+1}\lambda + g_{m-1,n}\lambda_1 + g_{m,n-1}\mu_1 + r \right].$$

If (m, n) is of type $2a$, this is the state where there is no existing limit order sellers. As illustrated in Figure 2.1, we have $m = 0$. Then set $f_{0,n} = \overline{P}$. The arrivals of patient seller, patient buyer and impatient seller might incur respectively activities

B, D and $C(1)$. One obtains the formula for $g_{0,n}$ as

$$g_{0,n} = \frac{1}{\Sigma_2}\left[g_{1,n}\mu + g_{0,n+1}\lambda + g_{0,n-1}\lambda_1 + r\right].$$

For type $2b$, as illustrated in Figure 2.1, we have $n = 0$. So the recursive equations under this situation, $f_{m,0}$ and $g_{m,0}$, are similar to those for type $2a$.

Next, we look at the boundary states when mixed strategies might happen. If state (m, n) is of type $3a$, set by convention $f_{0,n} = \overline{P}$. Since the capacity of limit order buyers reaches its maximum and there is no existing limit order sellers, event $E(2)$ is not possible but $C(3)$ might happen at price $\overline{P}$. The arrivals of new patient seller will place a limit order and stay in the book, which implies activity B not $A(1)$. The arrival of new impatient seller incurs activity $C(1)$. Thus, one get the formula for limit order sellers' utility

$$g_{0,n} = \frac{1}{\mu + \nu_{0,n} + \lambda_1}\left[g_{1,n}\mu + g_{0,n-1}\nu_{0,n} + g_{0,n-1}\lambda_1 + r\right].$$

If (m, n) is of type $4a$, existing limit order sellers reach its maximum capacity. Therefore the limit order sellers will not wait but trade following activity $E(1)$. If a new patient seller arrives, she will immediately submit a market sell order and trade with the top limit order buyer at $b_{m,n}$. This implies that activity $C(2)$ might happen. If a fnew patient buyer arrives, it is still better to wait than trade immediately for him. So event D is possible. The trading of impatient buyer and seller will incur event $A(1)$ and $C(1)$. One obtains the utility functions

$$\begin{aligned} f_{m,n} = \frac{1}{\Sigma_1 + \nu_{m,n}}[&f_{m,n-1}\mu + f_{m,n+1}\lambda + f_{m-1,n}\mu_1 \\ &+ f_{m,n-1}\lambda_1 - r + f_{m-1,n-1}\nu_{m,n}] \end{aligned}$$

and

$$\begin{aligned} g_{m,n} = \frac{1}{\Sigma_1 + \nu_{m,n}}[&g_{m,n-1}\mu + g_{m,n+1}\lambda + g_{m-1,n}\mu_1 \\ &+ g_{m,n-1}\lambda_1 + r + g_{m-1,n-1}\nu_{m,n}]. \end{aligned}$$

If (m, n) is of type 5, in the case of impatient traders arrival, they trade as normal with market orders, which are events B and D. In the case of patient traders arrival, the order book is in a state where it is better to trade immediately than wait. So patient traders will either do $E(1), E(2)$ or do $A(2), C(2)$. Thus, one obtains the formula

$$f_{m,n} = \frac{1}{\Sigma_1 + 2\nu_{m,n}}[f_{m,n-1}\mu + f_{m-1,n}\lambda + f_{m-1,n}\mu_1 \\ + f_{m,n-1}\lambda_1 - r + 2f_{m-1,n-1}\nu_{m,n}]$$

and

$$g_{m,n} = \frac{1}{\Sigma_1 + 2\nu_{m,n}}[g_{m,n-1}\mu + g_{m-1,n}\lambda + g_{m-1,n}\mu_1 \\ + g_{m,n-1}\lambda_1 + r + 2g_{m-1,n-1}\nu_{m,n}].$$

We can define $a_{m,n}$ and $b_{m,n}$ based on the ideas of the same expected utility in equilibrium as in the one-side case. More specifically, the one who is getting executed must have the same expected utility as the other traders in the book on the same side. They are given by the following formulas:

$$\begin{cases} a_{m,n} = f_{m-1,n} \text{ and } b_{m,n} = g_{m,n-1}, & \text{if } (m,n) \text{ is of type 1;} \\ a_{0,n} = \overline{P} \text{ and } b_{0,n} = g_{0,n-1}, & \text{if } (m,n) \text{ is of type 2a;} \\ a_{m,0} = f_{m-1,0} \text{ and } b_{m,0} = \underline{P}, & \text{if } (m,n) \text{ is of type 2b.} \end{cases}$$

For those points on the boundary, without loss of generality, we will give the derivation of $a_{m,n}$ at state of type 5. There are three situations where the existing bottom limit order seller gets executed: trade to a market buy order in cases of $A(1), A(2)$ or do mixed strategy $E(1)$. With probability $\frac{\mu_1+\lambda}{\mu_1+\lambda+\nu_{m,n}}$, the bottom seller trades to market buy order and receives $a_{m,n}$, while the other limit order sellers get $f_{m-1,n}$. With probability $\frac{\nu_{m,n}}{\mu_1+\lambda+\nu_{m,n}}$, the bottom seller switches to a limit order at some lower level h, and the top limit order buyer immediately switches to a market buy order and accepts the offer at h. This price h should be one such that for both the bottom seller and top buyer not losing utility than $f_{m,n}$ and $g_{m,n}$ respectively. So it has to be that $h = f_{m,n} = g_{m,n}$. All the other sellers now get utility $f_{m-1,n-1}$. The same expected utility for all sellers leads to the relationship

$$a_{m,n}(\mu_1 + \lambda) + \nu_{m,n} f_{m,n} = f_{m-1,n}(\mu_1 + \lambda) + \nu_{m,n} f_{m-1,n-1}.$$

Thereafter, we get the formulas

$$a_{m,n} = f_{m-1,n} + \frac{\nu_{m,n}}{\lambda + \mu_1}(f_{m-1,n-1} - f_{m,n})$$

and

$$b_{m,n} = g_{m,n-1} + \frac{\nu_{m,n}}{\lambda + \mu_1}(g_{m-1,n-1} - g_{m,n}).$$

The derivations for other types of boundary points are similar.

Theorem 3 in Rosu [74] gives the existence of a rigid, competitive Markovian equilibrium in the two-side case. We summarise it here: If (m, n) is in the interior of set Ω, the bottom seller places a limit sell order at $a_{m,n}$, and the top buyer places a limit buy order at $b_{m,n}$. If $(m, n) \in \gamma$, then the strategy is the same as the one above, except that with the first arrival in a Poisson process with intensity $(\nu_{m,n})$ the bottom seller changes the limit order from $a_{m,n}$ to $h = f_{m,n} = g_{m,n}$, and the top buyer immediately accepts it via a market buy order; the top buyer would not accept any higher limit sell order. If $(m, n) \notin \Omega$ and $m, n > 0$, then the bottom seller places a limit order at $h = f_{m,n} = g_{m,n}$ and the top buyer immediately accepts it via a market order. If $(m, n) \notin \Omega$ and $n = 0$, then the bottom seller places a market order at $\underline{P}$ and exists the game.

2.2 Price impact and resilience

We present some of our main contributions in this section, we firstly translate the definitions of the price impact and the price overshoot in Rosu [74] by notions of two-side resilience correspondingly. Secondly, for the equilibrium strategies in one-side case, we give a rigourous proof of the theorem 2 in Rosu [74] by introducing the proposition 2.2.2. Thirdly, we provide a rigourous proof for same side resilience in the Proposition 2.2.3. Finally, we also provide an counterexample of the conjecture in Rosu [74], which stated that the same side resilience exists without the restriction on arrival rates μ_i and μ.

2.2.1 Resilience measurement

Recall that the same side resilience is the decay of the price impact on the same side of the large market order. The opposite side resilience is the decay of the price impact on the opposite side of the large order.

When an $i-$unit market buy order is submitted to the order book of m limit order sellers, this market order clears the sell orders from the lowest one to the lowest ith order. The lowest $(i+1)$th limit sell order immediately becomes the new best ask order in the order book. This is the immediate price impact incurred by the large order. Since the agents are fully strategic, they instantly regroup to adjust to the new state with $m - i$ sellers. The second best ask price changes corresponds to the transit decay of the first price change.

In Rosu's model [74], these two price changes take place instantaneously, and are described by the notions of temporary and permanent impact. However, as stated in Chapter 1, the permanent price impact is observed in the long term time. To prevent misunderstanding, in this thesis the permanent price impact is the second price change and the temporary price impact is the first price change as discussed above. Without loss of generality, we will define the temporary and permanent price impact on the ask side of the order book. These definitions could be extended to the bid side in a similar way.

Definition 2.2.1: *Consider the limit order book with only patient sellers and impatient buyers who can submit a market order of size at most k-unit for some $k > 1$. In the state of m patient sellers, denote by $a_m(j)$ the ask price of the jth limit sell* *order starting from the best ask price* $a_m(1)$ *for* $j \le m$. *Denote by* $i_0 := \min\{k, m\}$. *The temporary price impact on the ask side caused by a buy market order of size* $i \le i_0$ *is defined as the difference*

$$I_m^A(i) := a_m(i+1) - a_m(1), \tag{2.2.1}$$

which is the difference between the $i + 1'st$ *offer* $a_m(i+1)$ *from the bottom and the best ask price* $a_m(1)$.

The permanent price impact on the ask side caused by a buy market order of size $i \le i_0$ *is given by*

$$P_m^A(i) := a_{m-i}(1) - a_m(1), \tag{2.2.2}$$

which is the difference between the best ask price $a_{m-i}(1)$ *in the state with* $m - i$ *sellers and the best ask price* $a_m(1)$ *before the market order was submitted.*

If use the superscription A for the ask side and B for the bid side, we say there is the same side resilience if the difference between same side temporary impact and permanent impact is positive, i.e. $I_m^A(i) - P_m^A(i) > 0$ and $I_m^B(i) - P_m^B(i) > 0$. The opposite side resilience is measured by the positivity of the opposite side, or non-execution side, permanent impact, i.e. $P_m^B(i) > 0$ if $I_m^A(i)$ is non-negative and $P_m^A(i) > 0$ if $I_m^B(i)$ is non-negative.

2.2.2 Solution of the recursive system in the one-side case

In this part, we will discuss the properties of the solution of the recursive system proposed in the one-side order book case. This is an important step in proving the same side resilience and opposite side resilience. In general, it is difficult to get an explicit solution of the coupled recursive system in the two-side case. Consider the limit order book with only patient sellers and impatient buyers who can submit up to k-unit market orders for some $k > 1$, placing orders between the price range $[\underline{P}, \overline{P}]$; patient sellers arrive at rate $\mu > 0$ and have waiting cost discount factor $r > 0$; impatient buyers who submit i-unit market orders arrive at rate $\mu_i > 0$ if $i \leq k$ and $\mu_i = 0$ if $i > k$. Let us recall the recursive equations of the sellers' expected utility f_m with the boundary conditions given by

$$\begin{cases} f_i = \overline{P} \text{ if } i \leq 0, \\ \left(1 + \frac{\Sigma}{\mu}\right) f_m = f_{m+1} + \sum_{i\geq 1} \frac{\mu_i}{\mu} f_{m-i} - \frac{r}{\mu} \text{ for } 1 \leq m \leq M-1, \\ \left(1 + \frac{\Sigma}{\nu}\right) f_M = f_{M-1} + \sum_{i\geq 1} \frac{\mu_i}{\nu} f_{M-i} - \frac{r}{\nu}, \\ f_M = \underline{P}, \end{cases} \tag{2.2.3}$$

where $\Sigma = \sum_{i\geq 1} \mu_i$ and $\mu_j = 0$ if $j > k$.

The structure of the solution for the recursive system (2.2.3) is presented in Proposition 2.2.2.

Proposition 2.2.2: *Given the difference equation*

$$\left(1 + \frac{\Sigma}{\mu}\right) f_m = f_{m+1} + \sum_{i\geq 1} \frac{\mu_i}{\mu} f_{m-i} - \frac{r}{\mu} \quad \textit{for} \quad 1 \leq m \leq M-1$$

with $\Sigma = \sum_{i\geq 1} \mu_i$ *and* $\mu_j = 0$ *if* $j > k$. *Assume that* $\mu > \sum_{i\geq 1} i\mu_i$. *The structure of the solution to this difference equation is given by*

$$f_m = C_0 + C_1\alpha_1(m) + \ldots + C_k\alpha_k(m) + \frac{r}{\mu - \sum_{i\geq 1} i\mu_i} m$$

where $\alpha_0 = 1$ *and* $\alpha_1(m), \ldots, \alpha_k(m)$ *are the corresponding ansatz solutions of the difference equation* (2.2.2). $\alpha_i(m)$ *are functions of* m *and* $\mu, \mu_1, ..., \mu_k$.

The constants $C_0, \ldots, C_k$ *and* M *are determined by the boundary conditions* $f_i = \overline{P}$ *if* $i \leq 0$, $f_M = \underline{P}$ *and the state* M *difference equation*

$$\left(1 + \frac{\Sigma}{\nu}\right) f_M = f_{M-1} + \sum_{i\geq 1} \frac{\mu_i}{\nu} f_{M-i} - \frac{r}{\nu}. \tag{2.2.4}$$

2.2.3 Same side resilience

As proposed in Assumption 1 and Proposition 8 by Rosu [74], the same side resilience is obtained under the assumption that the market order arrival rate of $i \geq 2$-unit is much smaller than the arrival rate of one-unit order. One possible reason behind this assumption is due to the large order splitting behaviour in the market. Traders usually like to use small size order to eliminate the adverse price change. Thus, it is natural to make this assumption and study the market under it. Besides, we will test the attempt to relax this assumption in the end of this section. It is shown there that if the arrival rates of impatient traders satisfy are all the same, there does not exists the same side resilience in this microstructure model.

Before providing a mathematically rigorous proof of the same side resilience, we base on and improve the notation in Rosu [74]. For $i \geq 2$, denote by

$$\mu_i = \epsilon \phi_i \mu_1 \text{ with } \epsilon_i \geq 0 \text{ and } 0 \leq \phi_i < \infty.$$

The numbers ϕ_i are called the relative arrival rates. The utility function f_m and the ask prices $a_m(i)$ are now functions of μ, μ_1, ϵ and $\phi_{i>1}$.

Recall that the same side resilience is measured by the positivity of the difference between same side temporary impact and permanent impact, i.e. $I_m^A(i) - P_m^A(i) > 0$. The presence of the same side resilience[6] is summarised in the following proposition.

Proposition 2.2.3: *Consider the limit order book with only patient sellers and impatient buyers who can submit up to k-unit market orders for some $k > 1$, placing orders between the price range $[\underline{P}, \overline{P}]$; patient sellers arrive at rate $\mu > 0$ and have waiting cost discount factor $r > 0$; impatient buyers who submit i-unit market orders arrive at rate $\mu_i > 0$ if $i \leq k$ and $\mu_i = 0$ if $i > k$. If the arrival rates of the impatient*

[6] The rational behind the same side resilience effect given by Rosu [74], is that in order to take advantage of larger incoming market orders, the patient sellers stay higher price in the book. Once an i-unit market order hits the order book, the rest of patient sellers readjust because of the (off-equilibrium) competitive behaviour.

buyers satisfy the following conditions:

1. $\mu > \sum_{i\geq 1} i\mu_i$, *i.e. patient sellers arrive faster than the units demanded by the impatient buyers,*

2. $\epsilon \to 0^+$, *i.e. the arrival rates* $\mu_{i>1}$ *are much smaller than* μ_1,

the temporary price impact $I_m^A(i)$ *is larger than the permanent price impact* $P_m^A(i)$.

A sketch of the proof: step 1, one proves the continuity of the ansatz solutions $(\alpha_0, \alpha_1, ..., \alpha_k) \in \mathbb{R}^{k+1}$ of the difference equation 2.2.2 at the point $(\mu, \mu_1, 0, ..., 0) \in \mathbb{R}^{k+1}$; step 2, one needs to find the limit $\lim_{\epsilon\to 0^+} f_m$ for any $m \leq M$; step 3, one proves the limit of the consecutive difference of f_m is positive; step 4, one estimates the asymptotic behaviour of the differences $a_{m-i}(1) - f_{m-i-1}$ and $a_m(i+1) - f_{m-i-1}$.

Conjectured in Rosu [74], he speculated that the same side resilience holds without the assumption on fast decaying arrival rates μ_i for $i \geq 2$. However, by the following counterexample, it is shown that this conjecture does not hold when the arrival rates μ_i are all the same for $i \geq 1$.

Proposition 2.2.4: *Consider the limit order book with only patient sellers and impatient buyers who can submit up to k-unit market orders for some* $k > 1$, *placing orders between the price range* $[\underline{P}, \overline{P}]$; *patient sellers arrive at rate* $\mu > 0$ *and have waiting cost discount factor* $r > 0$; *impatient buyers who submit i-unit market orders arrive at rate* $\mu_i > 0$ *if* $i \leq k$ *and* $\mu_i = 0$ *if* $i > k$.

If the arrival rates $\mu_i = \mu_{i+1}$ *for* $i = 1, ..., k-1$, *then there does not exist the same side resilience in this trading game.*

2.2.4 Opposite side resilience

Recalling that in the two side order book case, we only consider $k = 1$, i.e. only one-unit market orders are considered. The opposite side resilience is measured by the positivity of the opposite side permanent impact given there are temporary impact on the same side of the larger order.

In order to prove the opposite side resilience, Rosu [74] suggests to regard the limit order book as the collection of the ask side and bid side. The reservation value for each side is given by the best price on the other side. Without loss of generality in state $(m+1, n)$, the sell side of the book can be considered as a one-side model with lower bound of price range $\underline{P} = b_{m+1,n}$, i.e. the best bid price. When the best bid price $\underline{P}$ moves down to $\underline{P} - \Delta$ for some $\Delta > 0$, which is caused by an one-unit market order, one can observe a fall of the ask price $a_{m+1,n}$ too in this model. The Proposition 10 in Rosu [74] gives an approximation of the magnitude of the opposite side resilience. We review and prove it here again for consistency reason. From the proof, we should also note that this proposition works when the order book is not full, i.e. the interior of the state region Ω. This is due to an approximation expression of $a_{m+1,n}$ by $f_{m,n}$.

Proposition 2.2.5: *Suppose the limit order book is in the state with $m+1$ sellers and n buyers, where $(m+1, n)$ is not on the boundary set γ. Assume that $\lambda = \mu$, and $\lambda_1 = \mu_1$ satisfying $c = \frac{\mu}{\mu_1} > 1$ and $\lambda_i = \mu_i = 0$ for $i > 1$. Then if a market sell order moves the best bid price down by Δ, the best ask price moves down approximately by $\Delta(1 - \frac{1}{c^m})$. Therefore, there exists the opposite side resilience in this LOB model.*

2.3 Proofs

Proof of Proposition 2.2.2. The difference equation (2.2.2) can be rewritten as

$$(\mu + \Sigma) f_m + r = \mu f_{m+1} + \mu_1 f_{m-1} + \ldots + \mu_k f_{m-k}.$$

The corresponding homogeneous equation is

$$\mu f_{m+1} - (\mu + \Sigma) f_m + \mu_1 f_{m-1} + \ldots + \mu_k f_{m-k} = 0.$$

The auxiliary equation is given by

$$\mathcal{P}_{k+1}(x) = \mu x^{k+1} - (\mu + \Sigma)x^k + \mu_1 x^{k-1} + \ldots + \mu_k.$$

Since all μ_i are positive for $i = 1, ..., k$, the roots $\beta_0 = 1, ..., \beta_k$ of the auxiliary equation $\mathcal{P}_{k+1}(x)$ are all not zero. Considering the possibilities of complex roots and repeated roots, we denote by α_i with $i = 0, ..., k$ the ansatz solutions of the difference equation (2.2.2). The general solution of the homogeneous equation can be expressed as

$$f_m = C_0 + C_1\alpha_1 + \ldots + C_k\alpha_k.$$

Also it is easy to check that $\frac{r}{\mu - \sum_{i=1}^{k} i\mu_i} m$ is a special solution for the difference equation (2.2.2).

$\square$

Step 1: prove the continuity of $(1, a, \alpha_2, ..., \alpha_k)$ at the point $(\mu, \mu_1, 0, ..., 0)$

Recall that $\beta_0, \beta_1, ..., \beta_k$ are the roots of the auxiliary polynomial $\mathcal{P}_{k+1}(x)$ and $\alpha_0, \alpha_1, ..., \alpha_k$ are the corresponding ansatz solutions of the difference equation (2.2.2). We want to know the behaviour of α_i for $i = 1, 2, \cdots, k$ as $\epsilon \to 0^+$.

We will apply the continuity theorem of the roots of a polynomial given in Cucker and Corbalan [26].

Theorem 2.3.1: *Let $P_n(x) = x^n + a_1 x^{n-1} + \ldots + a_n$ be a monic complex polynomial, and let $\xi_1, \ldots, \xi_n$ be its roots. Given a real number $\omega > 0$, there is a real number $\delta > 0$ such that for every monic polynomial $Q_n(x) = x^n + b_1 x^{n-1} + \ldots + b_n$, if $|b_j - a_j| < \delta$ for $1 \leq j \leq n$, then there are $\zeta_1, \ldots, \zeta_n \in \mathbb{C}$ such that $Q_n(x) = \prod_{1 \leq j \leq n}(x - \zeta_j)$ and $|\zeta_j - \xi_j| < \omega$ for $1 \leq j \leq n$.*

Lemma 2.3.2: *Under the assumptions of Proposition 2.2.2, let $(\beta_0, ..., \beta_k)$ be the roots of the polynomial $\mathcal{P}_{k+1}(x) = \mu x^{k+1} - (\mu + \sum_{i=1}^{k} \mu_i)x^k + \sum_{i=1}^{k} \mu_i x^{k-i}$. Given the polynomial $\mathcal{Q}_{k+1}(x) = \mu x^{k+1} - (\mu + \mu_1)x^k + \mu_1 x^{k-1}$ with roots $(b_0, ..., b_k)$, if the* *condition (2) of Proposition 2.2.3 is satisfied, namely the arrival rate of more-than-one units orders $\mu_{i>1}$ are much smaller than arrival rate of one-unit order μ_1, then the roots $(\beta_0, ..., \beta_k)$ is continuous at the point $(\mu, \mu_1, 0, ..., 0)$, and as $\epsilon \to 0^+$ the roots $(\beta_0, ..., \beta_k)$ tends to $\left(1, \frac{\mu_1}{\mu}, 0, \ldots, 0\right)$ which is the roots of polynomial $\mathcal{Q}_{k+1}$.*

Moreover, the corresponding ansatz solutions $\alpha_0(m), ..., \alpha_k(m)$ of the difference equation (2.2.2) *are continuous at the point $(\mu, \mu_1, 0, ..., 0)$ and tend to the point $\left(1, \left(\frac{\mu_1}{\mu}\right)^m, 0, \ldots, 0\right)$.*

Proof. As $\epsilon \to 0^+$, the coefficients of polynomial $\mathcal{P}_{k+1}(x)$ approaches to the coefficients of polynomial $\mathcal{Q}_{k+1}(x)$. One can easily solve the equation $\mathcal{Q}_{k+1}(x) = 0$ and get the roots as $\left(1, \frac{\mu_1}{\mu}, 0, \ldots, 0\right)$. By Theorem 2.3.1, we immediately obtain the continuity of $(\beta_0, ..., \beta_k)$ at the point $(\mu, \mu_1, 0, ..., 0)$.

The guessed corresponding solutions α_i are power functions of β_i or linear combination of power functions of β_i. Thus, the $(\alpha_0, ..., \alpha_k)$ is continuous at the point $(\mu, \mu_1, 0, ..., 0)$. And for all $m \leq M$, one has $\lim_{\epsilon \to 0^+} \alpha_0(m) = 1$, $\lim_{\epsilon \to 0^+} \alpha_1(m) = \left(\frac{\mu_1}{\mu}\right)^m$ and $\lim_{\epsilon \to 0^+} \alpha_i(m) = 0$ for $i > 1$.

$\square$

Step 2: find the limit of f_m as $\epsilon \to 0^+$ for any $m \leq M$

Lemma 2.3.3: *Under the assumptions of Proposition 2.2.2, one has*

$$\lim_{\epsilon \to 0^+} f_m = \overline{P} + C\left(\left(\frac{\mu_1}{\mu}\right)^m - 1\right) + \frac{r}{\mu - \mu_1} m. \tag{2.3.1}$$

The constant C is given by

$$C = \frac{r}{\mu - \mu_1} \frac{\frac{\mu + \nu}{\mu_1 + \nu}}{\left(\frac{\mu_1}{\mu}\right)^{M-1} \left(1 - \frac{\mu_1}{\mu}\right)}.$$

Proof. For $j = 0, 1, \ldots, k-1$, the boundary conditions associated with difference equation (2.2.2) are

$$C_0 + C_1\alpha_1(-j) + \ldots + C_k\alpha_k(-j) = \overline{P} + j\frac{r}{\mu - \sum i\mu_i} \tag{2.3.2}$$

and

$$C_0 + C_1\alpha_1(M) + \ldots + C_k\alpha_k(M) = \underline{P} - M\frac{r}{\mu - \sum i\mu_i}.$$

As $\epsilon \to 0^+$, one gets $\alpha_i(j) \to 0$ for $i > 1$ by Lemma 2.3.2. So the limit $\alpha_i(-j) \to \infty$ for $i > 1$ as $\epsilon \to 0^+$ holds. We find that the limit of the RHS of equation (2.3.2) is finite as ϵ becomes very small. Then, for $i > 1$ it has to be $\lim_{\epsilon\to 0^+} C_i < \infty$ such that $C_i = o(\alpha_i(j))$. Thus, we get

$$\begin{aligned}
\lim_{\epsilon\to 0^+} f_m &= \lim_{\epsilon\to 0^+}\left(C_0 + C_1\alpha_1(m) + C_2\alpha_2(m) + \ldots + C_k\alpha_k(m) + \frac{r}{\mu - \sum i\mu_i}m\right) \\
&= C_0 + C_1 \lim_{\epsilon\to 0^+}\alpha_1(m) + \lim_{\epsilon\to 0^+} C_2 \lim_{\epsilon\to 0^+}\alpha_2(m) + \ldots + \lim_{\epsilon\to 0^+} C_k \lim_{\epsilon\to 0^+}\alpha_k(m) + \frac{r}{\mu - \mu_1}m \\
&= C_0 + C_1\left(\frac{\mu_1}{\mu}\right)^m + \frac{r}{\mu - \mu_1}m.
\end{aligned}$$

Since $C_i = o(\alpha_i(j))$ for $i > 1$, the asymptotic behaviour of the boundary conditions implies

$$\lim_{\epsilon\to 0^+} f_0 = \lim_{\epsilon\to 0^+}(C_0 + C_1 + \ldots + C_k) = C_0 + C_1 = \overline{P}$$

and

$$\lim_{\epsilon\to 0^+} f_M - \frac{r}{\mu - \mu_1}M = \lim_{\epsilon\to 0^+} C_0 + C_1\alpha_1(M) + \ldots + C_k\alpha_k(M) = C_0 + C_1\left(\frac{\mu_1}{\mu}\right)^M = \underline{P}.$$

We solve the above boundary conditions and get

$$C_1 = \frac{\overline{P} - \underline{P} + \frac{r}{\mu-\mu_1}M}{1 - \left(\frac{\mu_1}{\mu}\right)^M} < \infty \text{ and } C_0 = \frac{\underline{P} - \overline{P}\left(\frac{\mu_1}{\mu}\right)^M - \frac{r}{\mu-\mu_1}M}{1 - \left(\frac{\mu_1}{\mu}\right)^M} < \infty.$$

Thereafter, the limit of the utility function f_m can be rewritten as

$$f_m = \overline{P} + C\left(\left(\frac{\mu_1}{\mu}\right)^m - 1\right) + \frac{r}{\mu - \mu_1}m,$$

where the constant C is given by

$$C = \frac{r}{\mu - \mu_1} \frac{\frac{\mu+\nu}{\mu_1+\nu}}{\left(\frac{\mu_1}{\mu}\right)^{M-1}\left(1 - \frac{\mu_1}{\mu}\right)}.$$

$\square$

Step 3: find the consecutive difference of f_m.

Lemma 2.3.4: *Under the assumptions of Proposition 2.2.2, one obtains* $\lim_{\epsilon\to 0^+}(f_{m-1} - f_m) > 0$.

Proof. We estimate the consecutive difference of f_m as follows

$$\begin{aligned}
\lim_{\epsilon\to 0^+}(f_{m-1} - f_m) &= C\left(\frac{\mu_1}{\mu}\right)^{m-1}\left(1 - \frac{\mu_1}{\mu}\right) - \frac{r}{\mu - \mu_1} \\
&= \frac{r}{\mu - \mu_1} \frac{\frac{\mu+\nu}{\mu_1+\nu}}{\left(\frac{\mu_1}{\mu}\right)^{M-1}\left(1 - \frac{\mu_1}{\mu}\right)} \left(\frac{\mu_1}{\mu}\right)^{m-1}\left(1 - \frac{\mu_1}{\mu}\right) - \frac{r}{\mu - \mu_1} \\
&= \frac{r}{\mu - \mu_1}\left(\frac{\mu+\nu}{\mu_1+\nu}\left(\frac{\mu_1}{\mu}\right)^{m-M} - 1\right) \\
&> 0.
\end{aligned}$$

The last inequality holds as the assumption of $\mu > \mu_1$. $\square$

Denote by $L_{a,b} = \lim_{\epsilon\to 0^+}(f_a - f_b)$ for all $0 < a < b \leq M$. Since $f_a - f_b = (f_a - f_{a+1}) + (f_{a+1} - f_{a+2}) + \ldots + (f_{b-1} - f_b)$, we then get the general difference given by

$$\lim_{\epsilon\to 0^+}(f_a - f_b) = L_{a,a+1} + \ldots + L_{b-1,b} > 0.$$

Step 4: estimate the differences $a_{m-i}(1) - f_{m-i-1}$ and $a_m(i+1) - f_{m-i-1}$

Recall that the ask prices in equilibrium are given by $a_m(i) = \frac{\sum_{j\geq i}\mu_j f_{m-j}}{\sum_{j\geq i}\mu_j}$, and the arrival rates of market order for more than one unit are denoted by $\mu_i =$

$\epsilon\mu_1\phi_i$ with $\epsilon \geq 0$ and $0 \leq \phi_i < \infty$.

Lemma 2.3.5: *Under the assumption of Proposition 2.2.2, if the arrival rates $\mu_{i>1}$ are much smaller than μ_1, namely ϵ tends to the zero from right, the difference $a_{m-i}(1) - f_{m-i-1}$ is estimated as $0 < a_{m-i}(1) - f_{m-i-1} < o(1)$.*

Proof. According to Lemma 2.3.2 and Lemma 2.3.3, the difference $a_{m-i}(1) - f_{m-i-1}$ is continuous at the point $(\mu, \mu_1, 0, ..., 0)$. We can then compute the limit of the difference $a_{m-i}(1) - f_{m-i-1}$ as follows

$$
\begin{aligned}
&\lim_{\epsilon\to 0^+} (a_{m-i}(1) - f_{m-i-1}) \\
&= \lim_{\epsilon\to 0^+} \frac{\mu_1 f_{m-i-1} + \ldots + \epsilon\phi_k\mu_1 f_{m-i-k}}{\mu_1 + \epsilon\mu_1(\phi_2 + \ldots + \phi_k)} - \frac{\mu_1 f_{m-i-1} + \epsilon\mu_1(\phi_2 + \ldots + \phi_k) f_{m-i-1}}{\mu_1 + \epsilon\mu_1(\phi_2 + \ldots + \phi_k)} \\
&= \lim_{\epsilon\to 0^+} \frac{\epsilon}{1 + \epsilon(\phi_2 + \ldots + \phi_k)} (\phi_2(f_{m-i-2} - f_{m-i-1}) + \ldots + \phi_k(f_{m-i-k} - f_{m-i-1})) \\
&= \lim_{\epsilon\to 0^+} \frac{\epsilon}{1 + \epsilon(\phi_2 + \ldots + \phi_k)} \lim_{\epsilon\to 0^+} [\phi_2(f_{m-i-2} - f_{m-i-1}) + \ldots + \phi_k(f_{m-i-k} - f_{m-i-1})] \\
&= 0.
\end{aligned}
$$

The last equation holds because of Lemma 2.3.4. $\square$

Lemma 2.3.6: *Under the assumption of Proposition 2.2.2, if the arrival rates $\mu_{i>1}$ are much smaller than μ_1, one has $\lim_{\epsilon\to 0^+} (a_m(i+1) - f_{m-i-1}) > 0$.*

Proof. First, we note that for any $\epsilon > 0$, the difference is positive, i.e.

$$
\begin{aligned}
a_m(i+1) - f_{m-i-1} &= \frac{\phi_k f_{m-k} + \ldots + \phi_{i+1} f_{m-i-1}}{\phi_k + \ldots + \phi_{i+1}} - f_{m-i-1} \\
&= \frac{1}{\phi_k + \ldots + \phi_{i+1}} [\phi_k(f_{m-k} - f_{m-i-1}) + \ldots \\
&\qquad + \phi_{i+1}(f_{m-i-1} - f_{m-i-1})] > 0.
\end{aligned}
$$

Since f_m is decreasing with m, for $i+1 \leq j \leq k$, $m-k \leq m-j \leq m-i-1$ one has $f_{m-k} \geq f_{m-j} \geq f_{m-i-1}$. Thus the last inequality holds.

Next, we look at the asymptotic behaviour of the difference $a_m(i+1) - f_{m-i-1}$

when $\epsilon \to 0^+$. By Lemma 2.3.4, one gets

$$\begin{aligned}
&\lim_{\epsilon\to 0^+}(a_m(i+1) - f_{m-i-1}) \\
=& \lim_{\epsilon\to 0^+} \frac{1}{\phi_k + \ldots + \phi_{i+1}} [\phi_k(f_{m-k} - f_{m-i-1}) + \ldots + \phi_{i+1}(f_{m-i-1} - f_{m-i-1})] \\
=& \frac{1}{\phi_k + \ldots + \phi_{i+1}} [\phi_k L_{m-k,m-i-1} + \ldots + \phi_{i+1} L_{m-i-1,m-i-1}] > 0.
\end{aligned}$$

□

Step 5: to prove the same side resilience

Proof of Proposition 2.2.3. There is same side resilience if and only if $I^A_m(i) = a_m(i+1) - a_m(1) > a_{m-i}(1) - a_m(1) = P^A_m(i)$, which is equivalent to establish $a_m(i+1) > a_{m-i}(1)$.

By Lemma 2.3.5, we know $\lim_{\epsilon\to 0^+}(a_{m-i}(1) - f_{m-i-1}) = 0$. That implies that for $\forall\theta > 0$, there is $\delta > 0$ such that for all $0 < \mu_i < \delta$ with $i > 1$, one has $|a_{m-i}(1) - f_{m-i-1}| < \theta$. Furthermore, since the ask price $a_{m-i}(1)$ can be regarded as a weighted average of a decreasing sequence f_{m-i-k}, $f_{m-i-k+1}$,$\cdots$, f_{m-i-1}, one gets$a_{m-i}(1) > f_{m-i-1}$, which implies $a_{m-i}(1) < f_{m-i-1} + \theta$.

By Lemma 2.3.6, $\lim_{\epsilon\to 0^+}(a_m(i+1) - f_{m-i-1}) > 0$ implies that there are $\theta > 0$, for $\forall\delta > 0$, there are some μ_i with $i > 1$ satisfying $0 < \mu_i < \delta$ such that $a_m(i+1) > f_{m-i-1} + \theta$.

Therefore, the statement $I^A_m(i) > P^A_m(i)$ holds. □

Proof of Proposition 2.2.4. We prove by showing contradiction. Suppose $\mu_i = \mu_{i+1}$ for $i = 1, ..., k-1$. Substituting to the ask prices $a_m(i+1)$ and $a_{m-i}(1)$, we then get

$$a_m(i+1) = \frac{f_{m-k} + \ldots + f_{m-i-1}}{k-i}$$

and

$$a_{m-i}(1) = \frac{f_{m-i-k} + \cdots + f_{m-i-1}}{k}.$$

We want to see if the difference $a_m(i+1) > a_{m-i}(1)$ is positive. So we compute

$$
\begin{aligned}
& k(k-i)(a_m(i+1) - a_{m-i}(1)) \\
&= kf_{m-k} + ... + kf_{m-i-1} - (k-i)f_{m-i-k} - ... - (k-i)f_{m-k} - ... - (k-i)f_{m-i-1} \\
&= i(f_{m-k} + ... + f_{m-i-1}) - (k-i)[f_{m-i-k} + ... + f_{m-k+1}] \\
&= i\left[\left(1 - \frac{k}{i}\right)(f_{m-i-k} + ... + f_{m-k+1}) + f_{m-k} + ... + f_{m-i-1}\right] \\
&< i\left[f_{m-i-k}\left(1 - \frac{k}{i} + k - i\right) + \left(1 - \frac{k}{i}\right)(f_{m-i-k+1} + ... + f_{m-k+1})\right].
\end{aligned}
$$

When $i = 1$ the RHS of the inequality is negative. This is a contradiction to the same side resilience. □

Proof of Proposition 2.2.5. For any state $(m+1, n)$ in the interior of the state region Ω, one has the best ask price $a_{m+1,n}$ approximately equals $f_{m,n}$ with the dependence of f on n being omitted.

The temporary price impact of an one-unit market sell order is given by $I^B_{m+1}(1) = b_{m+1}(2) - b_{m+1}(1)$. As the best bid price (or the reservation price for the ask side traders) changes, the utility of limit order sellers changes by approximately $I^B_{m+1}(1) \times \frac{df_m}{d\underline{P}}$.

The dependence is computed as follows. By Proposition 2.2.2, for $k = 1$ and $m = M$ the expected utility function f_M is given by $f_M = \overline{P} + C(c^M - 1) + \frac{r}{\mu+\mu_1}M$, where $C = \frac{r}{\mu-\mu_1}\frac{\mu+\nu}{\mu_1+\nu}\frac{1}{c^{M-1}-c^M}$. At the same time by boundary condition $f_M = \underline{P}$, one has

$$\underline{P} = \overline{P} + C(c^M - 1) + \frac{r}{\mu + \mu_1}M. \tag{2.3.3}$$

Differentiating equation (2.3.3) with respect to $\underline{P}$, one gets

$$1 = \frac{dC}{d\underline{P}}(c^M - 1).$$

Then the expected utility function f_m depends on reservation value $\underline{P}$ in the following way:

$$\frac{df_m}{d\underline{P}} = \frac{1 - c^m}{1 - c^M}.$$

One can then approximate $\lim_{M\to\infty} \frac{df_m}{d\underline{P}} = 1 - \frac{1}{c^m}$. Thus, the approximated opposite side resilience is given by $\Delta\left(1 - \frac{1}{c^m}\right)$

$\square$

Chapter 3

The cross-impact LOB model and definitions of market irregularities

This chapter is our main contribution, which is the formulation of the *cross impact* LOB model. It is an extension of the LOB model of Obizhaeva and Wang [65]. The main difference and contribution compared to Obizhaeva and Wang [65] is that we include the same side resilience and opposite side resilience into the LOB market impact modelling framework and model a more general time-varying shape function. The model formulation is provided in Section 3.1.

Two existing LOB market impact models considered in Section 3.2.1 and Section 3.2.2 correspond to limiting cases of our cross-impact model developed in Section 3.1. We provide the derivations of these two existing models from our cross-impact model. Utilising the separation of the same side and opposite side resilience in our cross-impact model, a non-zero spread in a market impact model is achieved. In particular, the spread is an endogenous results of the two sides resilience effect instead of being exogenously given.

Finally, in Section 3.3, three market irregularity notions are provided. Two relationships about the absence of the three market irregularities are presented in Proposition 3.3.4 and Proposition 3.3.5, followed by two examples showing these hierarchy relationship are not invertible.

3.1 The model formulation

The optimal execution problem is characterised by the following standard assumptions in the LOB based transient market impact model framework. It is assumed there is only one large trader whose trading incurs price impact. All other traders are noise traders, whose trading activities determine the dynamics of the LOB when the large trader is not active. We do not presume the large trader has private information about the trading asset. There is some amount $Q > 0$ ($Q < 0$) of one asset to be bought (or sold) within a certain time period $[0, T]$ with $T < \infty$. The reason to trade is exogenously given outside of the optimal execution problem. The large trader wants to minimise the trading costs by splitting the large order Q into smaller pieces of market orders. We will use the superscription A and B to denote the 'ask' and 'bid' side for variables.

Trading strategy

In the rest of this thesis, we focus our effort on the set of deterministic execution strategies. Although it might be suboptimal in some circumstances according to Klöck [52] and Lorenz and Almgren [61], the deterministic strategy is a standard assumption and provides some very delightful insights on optimal execution problem and market irregularity issues, as shown in most of the market impact model literature where they focus on deterministic strategy as well. One can regard the deterministic strategy as sample paths of the stochastic strategy.

Define a trading strategy X to be composed of two non-decreasing processes X_t^A and X_t^B, which respectively represents the number of accumulative large buy and sell orders by large-order trader up to time $t \in [0, T]$. The trading strategy process X_t is defined via X_t^A and X_t^B as

$$X_t = \int_0^t dX_s^A - \int_0^t dX_s^B, \quad \text{with } X_0 = 0.$$

The admissible trading strategy set is defined by

$$\mathcal{A}(Q) := \big\{(X^A, X^B) : [0, T_+] \to \mathbb{R}_+^2 \;\big|\; X_0^i = 0,\ X_{T_+} = Q,\ X^i \text{ is non-decreasing} \\ \text{and bounded variation for } i = A, B\big\},$$

where $Q > 0$ ($Q < 0$) corresponds to a trading program of buying (selling) $|Q|$ shares of an asset.

In the case that trading takes place at discrete trading times $\mathcal{T} := \{t_0, t_1, ..., t_N\}$ with $0 \leq t_0 < t_1 < ... < t_N \leq T$, we constrain the admissible strategy set to

$$\mathcal{A}_N(Q) := \big\{(X^A, X^B) \in \mathcal{A}(Q) \;\big|\; X_t^i = X_{t_n+}^i \text{ a.e. on } (t_n, t_{n+1}] \text{ for } n = 0, 1, ..., N-1\big\} \\ \subset \mathcal{A}(Q).$$

Another subset of $\mathcal{A}(Q)$ which is important for the absence of market irregularity is the *pure* buy (sell) strategy set

$$\mathcal{A}_P(Q) := \big\{(X^A, X^B) \in \mathcal{A}(Q) \;\big|\; X_t^B = 0 \ (X_t^A = 0) \text{ a.e. } \forall t \in [0, T]\big\} \\ \subset \mathcal{A}(Q).$$

Both impulse trading and continuous trading are allowed in a trading strategy. For $i = A, B$, let us denote by dX_t^i the continuous strategy and by $\Delta X_t^i := X_{t+}^i - X_t^i$ the jumps of X^i at time t. Especially, we will denote the discrete buy (sell) order at time t_n by $\Theta_n^i := \Delta X_{t_n}^i$. We assume at a single point in time, only buy or sell order can be submitted, in the sense that $\Delta X_t^A \Delta X_t^B = 0$ and $dX_t^A dX_t^B = 0$ for all $t \in [0, T]$. Otherwise, only the net position is considered [1].

[1] If at time t, both buy and sell orders ΔX_t^A and ΔX_t^B are submitted by one large trader. Only a buy order of size $(\Delta X_t^A - \Delta X_t^B)\mathbb{1}_{\Delta X_t^A - \Delta X_t^B > 0}$, or a sell order of size $(\Delta X_t^B - \Delta X_t^A)\mathbb{1}_{\Delta X_t^B - \Delta X_t^A > 0}$ would be taken.

Best bid and ask price

In market impact models, it is a standard assumption that the actual trading price is decomposed into two parts, namely an unaffected price process which describes the price dynamics when the large trader is not trading and the price impact caused by the large trader. We denote by $(A_t^0)_{t\in[0,T]}$ the unaffected best ask price process and by $(B_t^0)_{t\in[0,T]}$ the unaffected best bid price process. They are exogenously assumed to be martingales on a given filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ satisfying $B_t^0 \leq A_t^0$.

The price impact caused by strategy (X^A, X^B) on the ask and bid side are denoted respectively by s^A and s^B. Therefore, when the large trader trades following strategy (X^A, X^B), the best ask price A_t and the best bid price B_t are defined as

$$A_t = A_t^0 + s_t^A(X_t^A, X_t^B),$$

and

$$B_t = B_t^0 - s_t^B(X_t^A, X_t^B).$$

Without loss of generality, we set $A_0^0 = A_0$ and $B_0^0 = B_0$. The dynamics of the price impact s^A and s^B will be given in the sequel.

Shape function

The shape function is an abstract description of the dynamics of limit order volumes at different prices. We will adopt the continuous shape function assumption as in most market impact models. The continuity assumption enables us to keep mathematical tractability and at the same time it makes good approximation of the real order book. The discrete tick size of real order book will result in the shape function being piece-wise constant. By approximation of the piece-wide constant step function with a smooth function, we can reach any degree of smoothness.

Assume that the ask side (bid side) shape function $f^A(t,x)$ $(f^B(t,x))$ is strictly positive, deterministic and twice continuous differentiable. The volume of limit sell and buy orders at time t between prices a and b is given respectively by $\int_a^b f^A(t,x)dx$ and $\int_a^b f^B(t,x)dx$.

Dynamics of price impact and volume impact

We introduce two sides resilience effect into the market impact modelling, which are the same side resilience and the opposite side resilience introduced in Section 1.2. We reinterpret them with notions of price impact in the following.

When a market buy order is matched against the limit sell orders on the ask side, the best ask price decreases. This reflects the price impact on the same side of the market order. We will denote by D_t^A the accumulative same side price impact caused by previous market buy orders up to time t and by D_t^B the accumulative same side price impact caused by previous market sell orders up to t. Set by convention $D_0^A = 0$ and $D_0^B = 0$.

Let us now turn to look at the evolution of the same side price impact. More specifically, one assumes that the same side price impact exponentially decays when the large trader is inactive. This exponential decay assumption is a natural and computational efficient one following the line of research of Obizhaeva and Wang [65] and Alfonsi et al. [5], though empirical studies suggested some other choices, such as power-law decay in Weber and Rosenow [85], Gatheral [32]. We denote by ρ_t the *same side resilience rate* for $t \in [0,T]$. It is assumed to be deterministic, positive and continuous differentiable. That is to say, given a trading strategy $(X^A, X^B) \in \mathcal{A}(Q)$, the same side price impact D^A recovers exponentially at rate ρ_t

$$dD_t^A = -\rho_t D_t^A dt + \frac{dX_t^A}{f(t, D_t^A)}, \tag{3.1.1}$$

and the same side price impact D^B recovers exponentially at rate ρ_t

$$dD_t^B = -\rho_t D_t^B dt + \frac{dX_t^B}{f(t, D_t^B)}, \tag{3.1.2}$$

where f^A and f^B are the strictly positive shape functions of the ask and bid side of the order book.

In addition, we know that a market buy order incurs price changes not only on the ask side, but also has an impact on the bid side. This reflects the price impact on the opposite side of the market order, and is called the *cross price impact*. Following the notation rule that superscription A is for variables on the ask side, and superscription B is for variables on the bid side, we denote by L_t^A the ask side cross price impact incurred by all previous market sell orders up to time t, and by L_t^B the bid side cross price impact incurred by all previous market buy orders up to time t. More precisely, they are functions of the form

$$L_t^A := L_t^A(X_t^B)$$

and

$$L_t^B := L_t^B(X_t^A),$$

with $L_0^A = 0$ and $L_0^B = 0$.

The creative part of this thesis is that we assume the size of cross price impact depends on the size of the same side price impact and exponentially decays by the *cross impact resilience rate* of $\beta > 0$. The dynamics of the cross price impact L^A and L^B are given as follows

$$dL_t^A = -(\beta + \rho_t)L_t^A dt + \beta D_t^B dt, \tag{3.1.3}$$

and

$$dL_t^B = -(\beta + \rho_t)L_t^B dt + \beta D_t^A dt. \tag{3.1.4}$$

Combining both the same side price impact and cross price impact into the best bid and ask price, one obtains the expressions for the price impact s^A and s^B

given by

$$s_t^A(X^A, X^B) = D_t^A(X^A) - L_t^A(X^B)$$

and

$$s_t^B(X^A, X^B) = D_t^B(X^B) - L_t^B(X^A)$$

with $s_0^A = s_0^B = 0$. In the sequel, to distinguish from the same side impact and cross impact, we will call s^A and s^B the *total price impact*.

As suggested in Alfonsi et al. [5] and Alfonsi and Acevdeo [2], there is another natural way to describe the market order execution and the transient decay in LOB based market impact model, namely it is assumed a *volume impact reversion*.

When a market buy order is submitted to the ask side, the existing limit sell orders are consumed by this market order. This reflects the *volume impact* on the same side of the market order. We denote by E_t^A the accumulative volume of limit sell orders consumed by all previous market buy orders up to time t and E_t^B the accumulative volume of limit buy orders consumed by all previous market sell orders up to time t. Set by convention that $E_0^A = 0$ and $E_0^B = 0$. We call E^A and E^B the same side volume impact.

It is assumed here that the same side volume impact decays exponentially by the rate of ρ_t when the large trader is not active. The same side resilience rate ρ_t is assumed to be deterministic, positive and continuous differentiable. Given a trading strategy $(X^A, X^B) \in \mathcal{A}(Q)$, the dynamics of the same side volume impact E^A and E^B are given by

$$dE_t^A = -\rho_t E_t^A dt + dX_t^A,$$

and

$$dE_t^B = -\rho_t E_t^B dt + dX_t^B.$$

At the same time, the opposite side volume impact is described by by J_t^A

and J_t^B. We call them the *cross volume impact*, which reflect the reactions on the opposite side of the market order in terms of volume of new limit sell and buy order. Following the same notation rule of the superscription A and B, we denote by J_t^A the ask side cross volume impact incurred by all previous market sell orders up to time t, and by J_t^B the bid side cross volume impact incurred by all previous market buy orders up to time t. Given a trading strategy (X^A, X^B), the dynamics of the cross volume impact follow the way

$$dJ_t^A = -(\beta + \rho_t)J_t^A dt + \beta E_t^B dt$$

and

$$dJ_t^B = -(\beta + \rho_t)J_t^B dt + \beta E_t^A dt$$

where β is the cross impact resilience rate and the initial conditions $J_0^A = 0$ and $J_0^B = 0$.

Combining both the same side volume impact and cross volume impact, one obtains the expressions for the total volume impact V^A and V^B

$$V_t^A(X^A, X^B) = E_t^A(X^A) - J_t^A(X^B)$$

and

$$V_t^B(X^A, X^B) = E_t^B(X^B) - J_t^B(X^A)$$

with $V_0^A = V_0^B = 0$.

Via the shape functions f^A and f^B, the relationship between total price impact and total volume impact can be expressed as

$$V_t^A = \int_0^{s_t^A} f^A(t,x)dx \tag{3.1.5a}$$

and

$$V_t^B = \int_{-s_t^B}^0 f^B(t,x)dx. \tag{3.1.5b}$$

An example of trading and resilience in discrete time are illustrated in Figure 3.1

and Figure 3.2. The strategy used is: buy size x_0 at time t_0 and wait until time t_1 at which buy another x_1 shares. Explanations of each trading activity is given under the plot.

Cost function

Now we are ready to look at the trading costs given a strategy $X \in \mathcal{A}(Q)$. The following two assumptions are helpful to simplify the expressions of the cost functions.

Assumption 3.1.1: *We assume $F^A(t,x)$ and $F^B(t,x)$ are functions such that $F^A(t,0) = F^B(t,0) = 0$ for all $t \in [0,T]$,*

$$F^A(t,x) = \int_0^x f^A(t,p)dp, \quad F^B(t,x) = \int_{-x}^0 f^B(t,p)dp,$$

and

$$\lim_{x\to\infty} F^A(t,x) = \infty, \quad \lim_{x\to\infty} F^B(t,x) = -\infty.$$

Denote the first derivative of $F^i(t,x)$ on t by $\eta^i(t,x)$, i.e. $\frac{\partial F^i}{\partial t}(t,x) = \eta^i(t,x)$ for $i = A, B$.

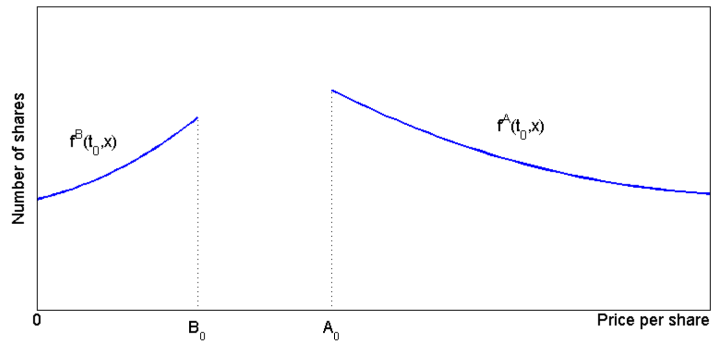
Assumption 3.1.2: *For each fixed time $t \in [0,T]$ and $i = A, B$, the function $F^i(t,x)$ is assumed to be invertible on x. Or equivalently, assume there exist functions $g^i(t,x)$ such that $F^i(t, g^i(t,x)) = x$.*

With the functions F^i and g^i, the relationship between the total price impact s^A, s^B and the total volume impact V^A, V^B can be rewritten as

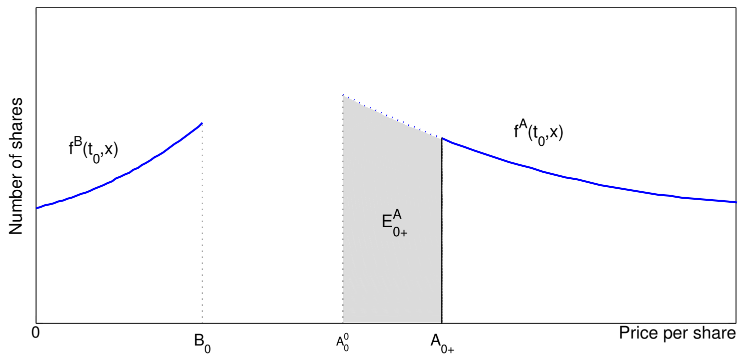
$$V_t^i = F^i(t, s_t^i) \text{ and } s_t^i = g^i(t, V_t^i). \tag{3.1.6a}$$

Before we derive the cost function, the following notations are needed. For $\forall x \in \mathbb{R}^+$, one takes

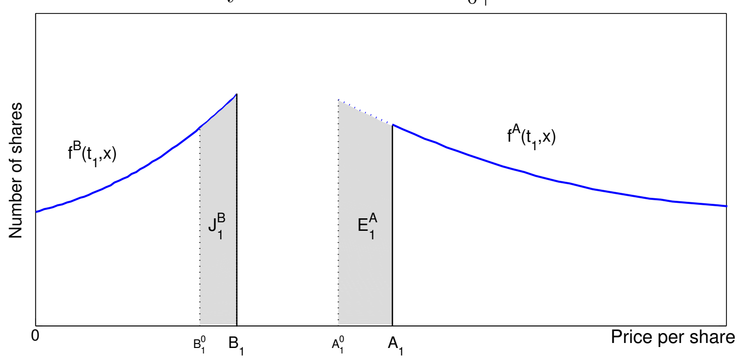
$$\tilde{F}^A(t,x) = \int_0^x y f^A(t,y)dy, \qquad G^A(t,x) = \tilde{F}^A(t, g^A(t,x)) \tag{3.1.7}$$



(a) From time 0 to t_0, the status of the order book without any large trader activity.

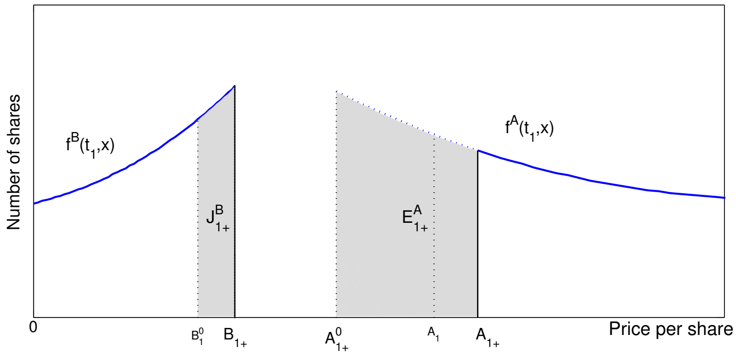


(b) At time t_{0+} a market buy order of size x_0 is executed against the existing limit sell orders on the ask side. The volume consumed is denoted by the shadow area E^A_{0+}.

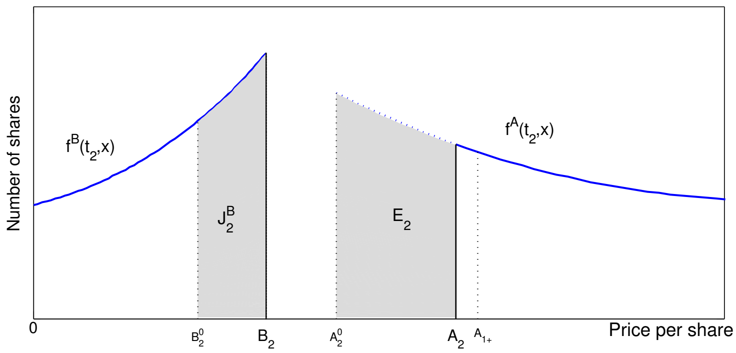


(c) Up to time t_1, there is no large trade during (t_0, t_1). Both sides of the order book are recovering. Reacted to market buy order x_0, the limit buy orders regenerate by volume of J^B_1 and the ask side volume impact decreases to E^A_1

Figure 3.1: LOB dynamics with strategy of buying x_0 at time t_0 and then waiting until t_1.



(a) At time t_{1+}, another buy order of size x_1 is placed. It consumes a volume of E^A_{1+} limit sell orders. The best ask prices jumps from A_1 to A_{1+}.



(b) During time (t_1, t_2), there is no large order trading. Both sides continue to recover. The bid side volume at time t_2 increases by J^B_2. The ask side recovers to new best ask price A_2.

Figure 3.2: Continued LOB dynamics with strategy of buying x_1 at time t_1 and then waiting until t_2.

and

$$\tilde{F}^B(t,x) = \int_{-x}^{0} y f^B(t,y)dy, \qquad G^B(t,x) = \tilde{F}^B(t, g^B(t,x)). \tag{3.1.8}$$

Furthermore, one can derive some regularity properties of $G^A(t,x)$ and $G^B(t,x)$ as summarised in the following.

Lemma 3.1.3: *For $i = A, B$, suppose $F^i(t,x)$ satisfies Assumption 3.1.1 and Assumption 3.1.2, the partial derivative on x of function $G^i(t,x)$ is given by $\frac{\partial G^i}{\partial x}(t,x) = g^i(t,x)$.*

The partial derivatives of $g^i(t,x)$ with $i = A, B$ are given as:

$$\frac{\partial g^i}{\partial t}(t,x) = -\left(\frac{\partial}{\partial g^i}\left(F^i(t,g^i) - x\right)\right)^{-1} \frac{\partial}{\partial t}(F^i(t,g^i) - x) = -\frac{\eta^i(t, g^i(t,x))}{f^i(t, g^i(t,x))}$$

and

$$\frac{\partial g^i}{\partial x}(t,x) = -\left(\frac{\partial}{\partial g^i}\left(F^i(t,g^i) - x\right)\right)^{-1} \frac{\partial}{\partial x}(F^i(t,g^i) - x) = \frac{1}{f^i(t, g^i(t,x))}.$$

Lemma 3.1.4: *For $i = A, B$, assume $g^i(t,x)$ is the inverse function of $F^i(t,x)$ such that $F^i(t, g^i(t,x)) = x$ and $F^i(t,x)$ satisfy Assumption 3.1.1 and Assumption 3.1.2. Then the functions $g^i(t,x)$ are strictly increasing with respect to x for all $x \in \mathbb{R}$ at some fixed time t.*

Since the best price takes the form of the sum of unaffected price and price impact, the expected trading cost $\mathcal{C}^\beta$ must be equal to the sum of the expected cost caused by the unaffected price and the expected cost caused by price impact. Let us first look at the expected costs caused by the unaffected price martingale. For a deterministic trading strategy $(X^A, X^B) \in \mathcal{A}(Q)$, using integration by parts to the expected cost function $\mathbb{E}\left[\int_0^T A_t^0 dX_t^A\right]$, one obtains

$$\begin{aligned}
\mathbb{E}\left[\int_0^T A_t^0 dX_t^A - B_t^0 dX_t^B\right] &= \mathbb{E}[A_{T+}^0 X_{T+}^A - A_0^0 X_0^A] - \mathbb{E}[B_{T+}^0 X_{T+}^B - B_0^0 X_0^B] \\
&= \mathbb{E}[A_{T+}^0 X_{T+}^A - B_{T+}^0 X_{T+}^B] = A_0^0 (X_{T+}^A - X_{T+}^B) = A_0^0 Q
\end{aligned}$$

since by definition we have $X_0^A = X_0^B = 0$, $X_{T+}^A - X_{T+}^B = Q$, $A_0^0 = B_0^0$ and A^0, B^0

are martingales.

Moreover, since only deterministic strategies are considered and the shape function, two sides resilience rates are all deterministic functions, the trading cost caused by price impact is actually deterministic. Therefore, in the rest of this thesis, without loss of generality we can set $A^0 = B^0 \equiv 0$ on time interval $[0,T]$ and denote by $\mathcal{C}^\beta$ the deterministic trading cost caused by price impact.

The costs of singular buy and sell trade ΔX_t^A and ΔX_t^B are given respectively by

$$\begin{aligned}\mathcal{C}^\beta(\Delta X_t^A) &= \int_{D_t^A - L_t^B}^{D_{t+}^A - L_t^B} x f^A(t,x)dx \\ &= \int_{g^A(t,E_t^A - J_t^B)}^{g^A(t,E_t^A + \Delta X_t^A - J_t^B)} x f^A(t,x)dx \\ &= G^A(t, E_t^A + \Delta X_t^A - J_t^B) - G^A(t, E_t^A - J_t^B)\end{aligned}$$

and

$$\begin{aligned}\mathcal{C}^\beta(\Delta X_t^B) &= \int_{-D_{t+}^B + L_t^A}^{-D_t^B + L_t^A} x f^B(t,x)dx \\ &= \int_{-g^B(t,E_t^B - \Delta X_t^B - J_t^A)}^{-g^B(t,E_t^B - J_t^A)} x f^B(t,x)dx \\ &= \left[G^B(t, E_t^B - \Delta X_t^B - J_t^A) - G^B(t, E_t^B - J_t^A)\right].\end{aligned}$$

The trading cost on $[0,T]$ of a continuous strategy dX^A is given by

$$\begin{aligned}\int_0^T \mathcal{C}^\beta(dX_t^A) &= \int_0^T \left[G^A(t, E_t^A - J_t^A + dX_t^A) - G^A(t, E_t^A - J_t^A)\right] \\ &= \int_0^T g^A(t, E_t^A - J_t^A)dX_t^A.\end{aligned}$$

The last equation holds because of Lemma 3.1.3. Similarly, we obtain the cost of continuous strategy dX^B

$$\int_0^T \mathcal{C}^\beta(dX_t^B) = \int_0^T G^B(t, E_t^B - J_t^b - dX_t^B) - G^B(t, E_t^B - J_t^b)$$

$$= \int_0^T g^B(t, E_t^B - J_t^B) dX_t^B.$$

For any deterministic trading strategy $X = (X^A, X^B) \in \mathcal{A}(Q)$, the trading cost function is then defined as

$$\begin{aligned} \mathcal{C}^\beta(X) = & \int_0^T g^A(t, V_t^A) dX_t^A + \sum_{t \leq T} \left[G^A(t, V_t^A + \Delta X_t^A) - G^A(t, V_t^A) \right] \\ & + \int_0^T g^B(t, V_t^B) dX_t^B + \sum_{t \leq T} \left[G^B(t, V_t^B - \Delta X_t^B) - G^B(t, V_t^B) \right]. \end{aligned} \tag{3.1.9}$$

Since both price impact reversion and volume impact reversion models are defined, it is convenient to write the cost function in terms of price impact s_t^i as well. It is given by

$$\begin{aligned} \mathcal{C}^\beta(X) = & \int_0^T s_t^A dX_t^A + \sum_{t \leq T} \left[G^A(t, F^A(t, s_t^A) + \Delta X_t^A) - G^A(t, F^A(t, s_t^A)) \right] \\ & + \int_0^T s_t^B dX_t^B + \sum_{t \leq T} \left[G^B(t, F^B(t, s_t^B) - \Delta X_t^B) - G^B(t, F^B(t, s_t^B)) \right]. \end{aligned} \tag{3.1.10}$$

3.2 Two limiting cases as the cross resilience rate $\beta \to \infty$ and $\beta \to 0$

In this section, we review two existing frameworks of LOB based market impact models and show that our cross-impact LOB model in Section 3.1 could reproduce their results by making $\beta \to \infty$ and $\beta \to 0$ respectively.

3.2.1 Zero-spread LOB model

Zero-spread assumption is widely used in different market impact models, like Almgren and Chriss [9], Alfonsi et al. [5], Gatheral et al. [37], Gatheral [32] and Alfonsi and Acevdeo [2]. The corresponding microstructure mechanics behind the zero-spread assumption is that once a market buy (sell) order is executed, the hole on the ask (bid)-side can be replenished by the incoming limit buy (sell) orders at an infinite speed. In terms of resilience, that is to say the opposite side resilience rate is infinite, i.e. $\beta \to \infty$. The price impact reversion in these zero-spread models can be reinterpret as a finite same side resilience rate, i.e. $\rho_t < \infty$ for all $t \in [0, T]$.

Now, we derive the cost function in the zero-spread LOB model by sending $\beta \to \infty$. As the cross impact resilience rate $\beta \to \infty$, one obtains the cross price impact on ask side L^A equals the same side price impact on bid side D^B, i.e. $L^A_t = D^B_t$. Similarly, one gets the relationships $L^B_t = D^A_t$, $J^A_t = E^B_t$, and $J^B_t = E^A_t$.

Denote by $s^{A,\infty}_t$, $s^{B,\infty}_t$ the total price impact in zero-spread framework and $V^{A,\infty}_t$, $V^{B,\infty}_t$ the total volume impact. They are expressed as

$$s^{A,\infty}_t = D^A_t - D^B_t = -s^{B,\infty}_t$$

and

$$V^{A,\infty}_t = E^A_t - E^B_t = -V^{B,\infty}_t.$$

Substituting $s^{A,\infty}$, $s^{B,\infty}$, $V^{A,\infty}$ and $V^{B,\infty}$ into the cost function (3.1.9), we then obtain

$$\begin{aligned} \mathcal{C}^{\infty}(X) = &\int_0^T s^{A,\infty}_t dX^A_t + \sum_{t \le T} \left[G^A(t, V^{A,\infty}_t + \Delta X^A_t) - G^A(t, V^{A,\infty}_t) \right] \\ &+ \int_0^T s^{B,\infty}_t dX^B_t + \sum_{t \le T} \left[G^B(t, V^{B,\infty}_t - \Delta X^B_t) - G^B(t, V^{B,\infty}_t) \right]. \end{aligned} \tag{3.2.1}$$

We will show how our cost function (3.2.1) coincides the zero-spread cost functions as given in Alfonsi et al. [5]. The key steps are to show $f^A(t,x) = f^B(t,x)$ and then $G^A(t,x) = G^B(t,-x)$. These can be proved by working on the relationship of $s^{A,\infty}_t = -s^{B,\infty}_t$ and $V^{A,\infty}_t = -V^{B,\infty}_t$. This is because equation

$$V^{A,\infty}_t = F^A(t, s^{A,\infty}_t) = -V^{B,\infty}_t = -F^B(t, s^{B,\infty}_t) = -F^B(t, -s^{A,\infty}_t)$$

implies equation $F^A(t,x) = -F^B(t,-x)$. Rewriting F^A and F^B via shape functions f^A and f^B, one has

$$\int_0^x f^A(t,p)dp = -\int_x^0 f^B(t,p)dp = \int_0^x f^B(t,p)dp.$$

Since both f^A and f^B are defined on the whole real line, the equality $f^A = f^B$ holds and also the equation $\tilde{F}^A(t,x) = \tilde{F}^B(t,x)$ holds. Furthermore, the relationship

$$s_t^{A,\infty} = g^A(t, V_t^{A,\infty}) = -s_t^{B,\infty} = g^B(t, -V_t^{A,\infty})$$

implies $g^A(t,x) = g^B(t,-x)$.

Recall the notation for $i = A, B$ one has $G^i(t,x) := \tilde{F}^i(t, g^i(t,x))$. The equation $g^A(t,x) = g^B(t,-x)$ implies the equation

$$\begin{aligned} G^A(t,x) &= \tilde{F}^A(t, g^A(t,x)) = \tilde{F}^A(t, g^B(t,-x)) \\ &= \tilde{F}^B(t, g^B(t,-x)) = G^B(t,-x) \\ &=: G(t,x). \end{aligned}$$

We can define a function $F(t,x)$ in a similar way as in cross-impact LOB model case. Set the shape function $f(t,x) := f^A(t,y)$.

Assumption 3.2.1: *Define function $F(t,x)$ by $F(t,x) := \int_0^x f(t,y)dy$, and assume that $F(t,0) = 0$ for all $t \in [0,T]$,*

$$\lim_{x \to -\infty} F(t,x) = -\infty$$

and

$$\lim_{x \to \infty} F(t,x) = \infty.$$

Its first derivative on t is denoted by $\eta(t,x)$, i.e. $\frac{\partial F}{\partial t}(t,x) = \eta(t,x)$.

Assumption 3.2.2: *For each fixed time $t \in [0,T]$, the function $F(t,x)$ is assumed to be invertible on x. Or equivalently, assume there is a function $g(t,x)$ such that*

$F(t, g(t,x)) = x$.

Remark 3.2.3 (Special cases of F and g)**:** *1) Zero-spread time independent shape function as in Alfonsi and Schied [3]: $f(t,x) := f(x)$. The anti-derivative function is $F(t,x) := F(x) = \int_0^x f(y)dy$ and it inverse function is $g(t,x) := F^{-1}(x)$.*

2) Zero-spread time-varying separable shape function as in Alfonsi and Acevdeo [2]: $f(t,x) := q(t)h(x)$. The corresponding F and g are $F(t,x) = q(t)F(x)$ where $F(x) = \int_0^x f(y)dy$ and $g(t,x) = F^{-1}\left(\frac{x}{q(t)}\right)$.

The admissible set of trading strategies under the zero-spread LOB model is defined as

$$\mathcal{A}^\infty(Q) := \Big\{ X_t : [0, T_+] \to \mathbb{R} \;\Big|\; X_t = \int_0^t dX_t^A - \int_0^t dX_t^B \\ \forall (X_t^A, X_t^B) \in \mathcal{A}(Q),\ X_0 = 0,\ \text{and } X_{T_+} = Q \Big\}.$$

The zero-spread discrete admissible strategy set can be defined as

$$\mathcal{A}_N^\infty(Q) := \big\{ X_t \in \mathcal{A}^\infty(Q) \;\big|\; X_t = X_{t_n+} \text{ on } (t_n, t_{n+1}] \text{ for } n = 0, 1, ..., N \big\}.$$

Denote by $\xi_t := \Delta X_t$ and short by $\xi_n := \xi_{t_n}$. The zero-spread cost function (3.2.1) of strategy $X \in \mathcal{A}^\infty(Q)$ can be rewritten as

$$\mathcal{C}^\infty(X) = \int_0^T (D_t^A - D_t^B) dX_t + \sum_{t \in [0,T]} \left[G(t, E_t^A - E_t^B + \Delta X_t) - G(t, E_t^A - E_t^B) \right]. \tag{3.2.2}$$

3.2.2 One-side LOB model

We call the LOB model proposed in Fruth et al. [31] by the *one-side* LOB model. By 'one-side', we mean there is no opposite side resilience modelled, namely $\beta = 0$. In the following, we remind the readers briefly the model of Fruth et al. [31].

Readers might have noticed that in Fruth et al. [31], at the very beginning both bid and ask sides of the order book were modelled. More specifically translated into our notation, the equations (1) and (2) in Fruth et al. [31] implies that for any $t \in [0,T]$ the best ask and best bid prices are

$$A_t = A_t^0 + s_t^A$$

and

$$B_t = B_t^0 - s_t^B,$$

where

$$s_t^A = \gamma X_t^A + D_t^A - \gamma \int_0^t \left(1 - e^{-\int_s^t \rho_u du}\right) dX_s^B$$

and

$$s_t^B = \gamma X_t^B + D_t^B - \gamma \int_0^t \left(1 - e^{-\int_s^t \rho_u du}\right) dX_s^A,$$

for the permanent impact factor γ. The price impact s_t^A, s_t^B are determined by both buy and sell trades via the permanent impact factor γ. However later in their Proposition 3.3, they claim only temporary impact should be considered. So they set the permanent impact factor $\gamma = 0$ for the rest of the analysis. By doing this, their LOB model only includes the same side resilience effect. Thus, in this thesis we call it one-side LOB model. However we should note that, when it is restricted that only buy (sell) order can be traded, there is no difference between the cross-impact LOB model and one-side LOB model.

Now let us derive the cost function in the one-side LOB model. As the cross impact resilience rate $\beta = 0$, one gets the cross price impact $L_t^A = 0$, $L_t^B = 0$ and the cross volume impact $J_t^A = 0$, $J_t^B = 0$ for all $t \in [0,T]$. Denote the total price impact in one-side LOB model by $s_t^{A,0}$, $s_t^{B,0}$, which are given by $s_t^{A,0} = D_t^A$, $s_t^{B,0} = D_t^B$.

Simply substituting $s_t^{A,0}$ and $s_t^{B,0}$ into cost function formula (3.1.10), we obtain the cost function of the one-side LOB model as

$$\begin{aligned}\mathcal{C}^0(X) = &\int_0^T D_t^A dX_t^A + \sum_{t\le T}\left[G^A(t, E_t^A + \Delta X_t^A) - G^A(t, E_t^A)\right] \\ &+ \int_0^T D_t^B dX_t^B + \sum_{t\le T}\left[G^B(t, E_t^B - \Delta X_t^B) - G^B(t, E_t^B)\right]. \end{aligned} \tag{3.2.3}$$

This expression coincides with the cost function in Fruth et al. [31].

3.3 Market irregularity definitions

In this section, we introduce three widely studied market irregularity notions in market impact model and some hierarchy relationships between them. More studies on the presence and absence conditions of these market irregularities will be presented in Chapter 5.

Note that as stated in Section 3.1, the effect of the unaffected price on the trading costs is not considered. We will recapitulate the original definition in terms of our notions. We drop the superscription ∞, β and 0 of the cost function unless stated otherwise, so that these definitions are not model dependent.

Definition 3.3.1: *A market impact model does not admit price manipulation (PMS) strategy if*

$$\inf_{X\in\mathcal{A}(0)} \mathcal{C}(X) \ge 0,$$

where $\mathcal{A}(0)$ is the set of the round trip strategies which means the total amount to trade is zero, i.e. $X_{T+} = 0$.

The notion of price manipulation strategy is first proposed in Huberman and Stanzl [48]. According to Gatheral and Schied [35], an optimal solution of the optimal execution problem does not exist if it is profitable to exploit some PMS strategy to a given strategy.

Definition 3.3.2: *A market impact model has positive liquidation costs (PLC) if for $\forall Q \in \mathbb{R}$, and every corresponding order execution strategy*

$$\inf_{X \in \mathcal{A}(Q)} \mathcal{C}(X) \geq 0,$$

where $\mathcal{A}(Q)$ is the admissible set of trading strategies that to get a total amount of Q shares.

Note that for round trips, PLC and absence of PMS are equivalent. However, PLC is defined on a bigger set of execution strategies.

Positive liquidation cost is first defined in Klöck et al. [53]. In its original paper, the condition is presented as $\inf_{X \in \mathcal{A}(Q)} \mathcal{C}(X) + A_0 Q \geq A_0 Q$ over all $X \in \mathcal{A}(Q)$ given $A_0 = B_0$. This PLC condition rules out the situation that on average the trader can make a profit beyond the face value $A_0 Q$ of a position out of the price impact incurred by his own trades.

Definition 3.3.3: *A market model does not admit transaction-triggered price manipulation (TTPM) strategy if for any $Q \in \mathbb{R}$,*

$$\inf_{X \in \mathcal{A}(Q)} \mathcal{C}(X) = \inf_{X \in \mathcal{A}_P(Q)} \mathcal{C}(X),$$

where $\mathcal{A}_P(Q)$ is the set of the pure trading strategies.

The TTPM strategy is first found in Alfonsi et al. [6]. It describes the situation in which the execution costs of a a buy (resp. sell) program can be reduced by intermediate sell (resp. buy) orders. For those models which admit TTPM strategy, optimal execution strategies may oscillate strongly between buy and sell orders, which implies the instability of the market impact models.

As preparation for the study of absence and presence condition of market irregularity in Chapter 5, we start with some hierarchy relationships between these three market irregularity definitions followed by some examples showing these relationships are not invertible.

Proposition 3.3.4: *In cross-impact LOB model with cost function (3.1.9), the absence of the transaction-triggered price manipulation implies the model has positive liquidation cost. Moreover, the existence of positive liquidation cost implies the absence of price manipulation strategies.*

We note that in the proof of the hierarchy relationship of market irregularity, a specific LOB model is required. Suppose given a general LOB model, the hierarchy relationship turns to be the one given in Proposition 3.3.5.

Proposition 3.3.5: *The absence of the transaction-triggered price manipulation implies the absence of price manipulation strategy.*

In the following, we will present some examples and show the above hierarchy relationships are not invertible.

Example 3.3.6 (No PMS but not PLC)**:** *By this example we will demonstrate a constant time-varying zero-spread LOB model that excludes the PMS but does not have the PLC under some circumstances.*

Assume the shape function is of the form $f(t,p) = q(t)f(x)$ *where we set* $f(x) = 1$ *and* $q(t) = 1 - bt + at^2$*. Let trading times be* $\mathcal{T} = \{0, 1, 2\}$*. For the trading strategy* $(x, y, Q - x - y)$*, we have the trading cost*

$$\begin{aligned}\mathcal{C}^\infty =& \frac{x^2}{2q(0)} + \frac{y}{2}\left(\frac{y}{2} + 2e^{-\rho}\frac{x}{q(0)}\right) \\ &+ \frac{Q-x-y}{2}\left(\frac{Q-x-y}{q(2)} + 2e^{-2\rho}\frac{x}{q(0)} + 2e^{-\rho}\frac{y}{q(1)}\right).\end{aligned}$$

Using the software Mathematica, one obtains $\mathcal{C}^\infty \geq 0$ *for* $Q = 0$ *under some specific combinations of the coefficients* a, b, ρ*. One example is taking* $a = 21.5$ *and* $b = -19.5$*,* $\rho = \log(2)$*.*

However, with the same coefficients a, b, ρ *and for any* $Q \neq 0$*, one obtains* $\inf \mathcal{C}^\infty \to -\infty$*, which implies the violation of the definition of positive liquidation cost.*

Example 3.3.7 (PLC but TTPM)**:** *We will present an example of a zero-spread*

LOB model which has PLC but admits the TTPM. Consider a purchase program to buy 50 *shares of some stock. Suppose that the resilience rate is* $\rho = 1$*, cross impact resilience rate is* $\beta \to \infty$ *and trading times* $\mathcal{T} = \{t_0 = 0, t_1 = 0.05, t_2 = 0.1\}$*, the shape function is of the form* $f(t,x) = q(t)$ *where* $q(t) = 4 + \cos(2\pi t)$*. We fix the first trading size to be* $\xi_0 = 20$*.*

As we change the size of the second trade ξ_1 *and the final trade* ξ_2 *accordingly such that a total size of* $Q = 50$ *is bought in the end, the change of the trading cost of strategy* $\{\xi_0, \xi_1, \xi_2\}$ *is shown in Figure 3.3. As we can observe, the optimal strategy* ξ_1 *is negative in some cases (shown by the negative part of* ξ_1*). In particularly, the minimum of the trading cost is obtained when* $\xi_1 = -5$ *and accordingly* $\xi_2 = 50 - \xi_1 - \xi_2 = 35$*. As a result, this model admits the TTPM. But this model has PLC by Proposition 5.3.2 in Chapter 5.*

3.4 Proofs

Proof of Lemma 3.1.3. Using the integration by parts formula $xF(x) = \int_0^x ydF(y) + \int_0^x F(y)dy$, for each fixed time $t \in [0,T]$ we could write

$$\begin{aligned}
G^A(t,x) &= \int_0^{g^A(t,x)} ydF^A(t,y) \\
&= g^A(t,x)F^A(t,g^A(t,x)) - \int_0^{g^A(t,x)} F^A(t,y)dy \\
&= \int_0^{g^A(t,x)} \big(x - F^A(t,y)\big)dy \\
&= \int_0^{g^A(t,x)} \int_{F^A(t,y)}^{x} dzdy \\
&= \int_0^{x} \int_0^{g^A(t,z)} dydz
\end{aligned}$$

and

$$\begin{aligned}
G^B(t,x) &= \int_{-g^B(t,x)}^{0} ydF^B(t,y) \\
&= g^B(t,x)F^B(t,g^B(t,x)) - \int_{-g^B(t,x)}^{0} F^B(t,y)dy
\end{aligned}$$

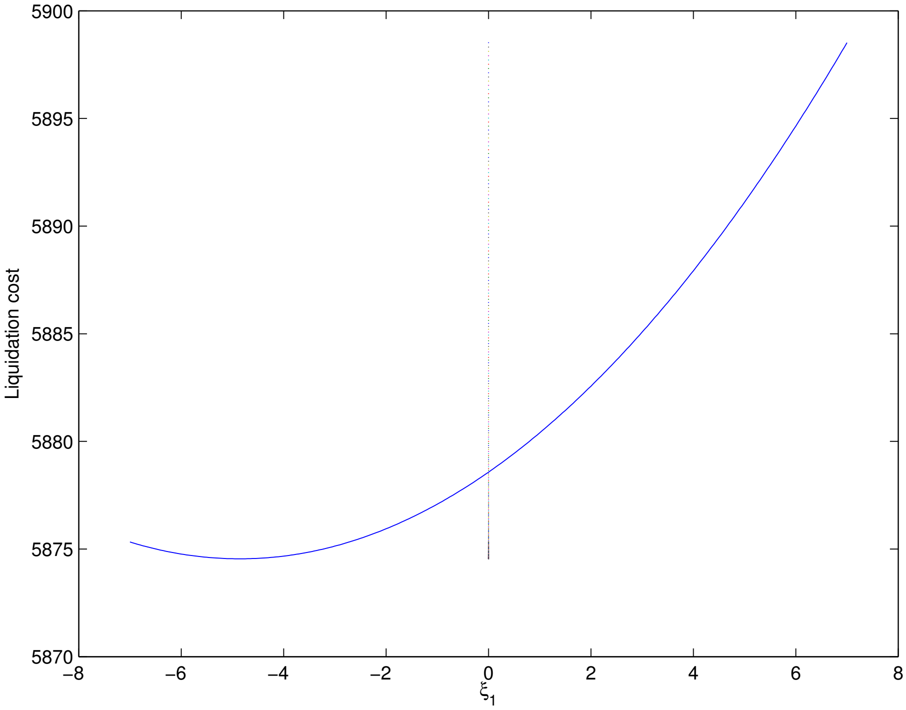


Figure 3.3: The change of the liquidation cost against the trade size ξ_1 at time t_1. The vertical line is $\xi_1 = 0$.

$$
\begin{aligned}
&= \int_{-g^B(t,x)}^{0} (x - F^B(t,y))dy \\
&= \int_{-g^B(t,x)}^{0} \int_{F^B(t,y)}^{x} dzdy \\
&= \int_{0}^{x} \int_{-g^B(t,z)}^{0} dydz.
\end{aligned}
$$

The last step in equations of G^A and G^B are both true because of the fact that the symmetric difference of the sets

$$
\begin{aligned}
&\big\{(z,y) \bigm| z \in [F^A(t,y), x] \text{ and } \in [0, g^A(t,x)]\big\}, \\
&\big\{(z,y) \bigm| z \in [0,x] \text{ and } y \in [0, g^A(t,z)]\big\}, \\
&\big\{(z,y) \bigm| z \in [F^B(t,y), x] \text{ and } y \in [-g^B(t,x), 0]\big\}
\end{aligned}
$$

and

$$
\big\{(z,y) \bigm| z \in [0,x] \text{ and } y \in [-g^B(t,z), 0]\big\}
$$

are both at most a countable union of line segments. Thus the two-dimensional Lebesgue measure is zero. Therefore, $G^i(t,x) = \int_0^x g^i(t,z)dz$ and $\frac{\partial G^i}{\partial x}(t,x) = g^i(t,x)$. $\square$

Proof of Lemma 3.1.4. **Ask side**: Suppose $x_2 < x_1 \leq \infty$. For some fixed time $t \in [0,T]$, choose y_1, y_2 such that $F^A(t,y_1) - x_1 = 0$ and $F^A(t,y_2) - x_2 = 0$. Thus we get $F^A(t,y_1) > F^A(t,y_2)$ since $F^A(t,x)$ is non-decreasing on x by Assumption 3.1.1. The inequality $F^A(t,y_1) > F^A(t,y_2)$ can be rewritten as $\int_0^{y_1} f^A(t,x)dx > \int_0^{y_2} f^A(t,x)dx$. The positivity of the shape function $f^A(t,x)$ implies $y_1 > y_2$, i.e. $g^A(t,x_1) > g^A(t,x_2)$.

Bid side: For $x_2 < x_1 \leq \infty$. Choose y_1, y_2 such that $F^B(t,y_1) - x_1 = 0$ and $F^B(t,y_2) - x_2 = 0$. Thus we have $F^B(t,y_1) > F^B(t,y_2)$, i.e. $\int_{-y_1}^0 f^B(t,x)dx > \int_{-y_2}^0 f^B(t,x)dx$. By the positivity of the shape function, we obtain $y_1 > y_2$, i.e. $g^B(t,x_1) > g^B(t,x_2)$. $\square$

Proof of Proposition 3.3.4. **Part One:** We prove the proposition under the price impact reversion and volume impact reversion respectively. Without loss of generality, we will consider a pure buy strategy $(X^A, 0) \in \mathcal{A}_P(Q)$. Recall the cost function of a cross-impact LOB model for pure strategy

$$\mathcal{C}^\beta(X_t^A, 0) = \int_0^T s^A dX_t^A + \sum_{t\in[0,T]} \left[G^A(t, V_t^A + \Delta X_t^A) - G^A(t, V_t^A)\right] \geq 0.$$

Case of price impact reversion: The total price impacts caused by strategy $(X^A, 0)$ are $s_t^A > 0$ and $s_t^B = 0$ for all $t \in [0, T]$. This is because the cross price impact $L^A(X^B)$ of a pure buy strategy is zero, namely $s_t^A = D_t^A$ for all $t \in [0, T]$. Considering the initial condition $D_0^A = 0$, the dynamics (3.1.1) and and X^A is non-decreasing, one gets $s_t^A > 0$. Via the relationship equation (3.1.6), the corresponding volume impact are given by $V_t^A = F^A(t, s_t^A) > 0$ and $V_t^B = 0$.

By Lemma 3.1.3 and Lemma 3.1.4, we know $G^A(t, x)$ is non-decreasing on x for some fixed $t \in [0, T]$. Therefore, for $\Delta X_t^A > 0$ and $V_t^A > 0$, $G^A(t, V_t^A + \Delta X_t^A) - G^A(t, V_t^A) \geq 0$ for $t \in [0, T]$. In addition, since dX_t^A is a positive measure on $[0, T]$, one obtains the first term in cost function is positive, i.e. $\int_0^T s_t^A dX_t^A \geq 0$.

The trading cost of arbitrary pure buy strategy $(X^A, 0)$ is then positive. So for any pure strategy one obtains $\inf_{X\in\mathcal{A}_P(Q)} \mathcal{C}^\beta(X) \geq 0$.

The absence of TTPM implies that $\inf_{X\in\mathcal{A}(Q)} \mathcal{C}^\beta(X) = \inf_{X\in\mathcal{A}_P(Q)} \mathcal{C}^\beta(X) \geq 0$. Therefore, the absence of TTPM implies the positive liquidation costs.

Case of volume impact reversion: The total volume impacts are $V_t^A = \int_0^t e^{-\rho(t-r)} dX_r^A > 0$ and $V_t^B = 0$ since X_A is an increasing process. By the relationship equation (3.1.6), the corresponding price impact are given by $s_t^A = g^A(t, V_t^A) > 0$ and $s_t^B = 0$ for all $t \in [0, T]$ since $g^A(t, x)$ is increasing on x.

Therefore, for $\Delta X_t^A > 0$ and $V_t^A > 0$, one has $G^A(t, V_t^A + \Delta X_t^A) - G^A(t, V_t^A) \geq 0$ for $t \in [0, T]$. Furthermore, dX_t^A is a positive measure on $[0, T]$, which implies $\int_0^T s_t^A dX_t^A \geq 0$.

As a result, the trading cost of an arbitrary pure buy strategy is positive, namely $\mathcal{C}^{\beta}(X_t^A, 0) \geq 0$. So for any pure strategy we have $\inf_{X \in \mathcal{A}_P(Q)} \mathcal{C}^{\beta}(X) \geq 0$. In particular, by absence of the TTPM strategy in cross-impact model, one obtains $\inf_{X \in \mathcal{A}(Q)} \mathcal{C}^{\beta}(X) = \inf_{X \in \mathcal{A}_P(Q)} \mathcal{C}^{\beta}(X) \geq 0$. Therefore, the absence of TTPM strategy implies the positive liquidation costs.

Part Two: For a round trip i.e. $Q = 0$, the model admits positive liquidation cost implies

$$\inf_{X \in \mathcal{A}(0)} \mathcal{C}^{\beta}(X) \geq 0,$$

which is just the definition of absence of price manipulation strategy. $\square$

Proof of Proposition 3.3.5. In this proof, we drop the superscription β, ∞ and 0 of the cost functions. By doing this, we show that the following arguments hold under all three models.

If we could prove that the existence of price manipulation strategy leads to transaction-triggered price manipulation, then the assertion holds.

On one hand, the existence of PM strategies implies that there is at least a round trip $X \in \mathcal{A}(0)$ such that $\mathcal{C}(X) < 0$, namely $\inf_{X \in \mathcal{A}(0)} \mathcal{C}(X) < 0$.

On the other hand, the absence of TTPM strategies implies that for $Q = 0$ one has $\inf_{X \in \mathcal{A}(0)} \mathcal{C}(X) = \inf_{X \in \mathcal{A}_P(0)} \mathcal{C}(X)$. While for a round trip, the pure strategy could only take the form of $X_t \equiv 0$ for all $t \in [0, T]$. As a result, the cost of a trading strategy $X \in \mathcal{A}_P(0)$ is always zero.

If the TTPM strategies are excluded, one has $\inf_{X \in \mathcal{A}(0)} \mathcal{C}(X) = \inf_{X \in \mathcal{A}_P(0)} \mathcal{C}(X) = 0$. This is a contradiction to the existence of PMS which states $\inf_{X \in \mathcal{A}(0)} \mathcal{C}(X) < 0$. Thus the TTPM strategy can not be excluded. Then the assertion holds. $\square$

Chapter 4

Application I: Optimal market order trading strategy

As a natural application of the cross impact LOB model, in Section 4.1 we will solve the optimal execution problem in a more general context. The complexity brought by modelling the two sides resilience ρ, β and a general time-dependent shape function $f(t,x)$ makes it hard to prove the existence of optimal strategy. Exploiting the Proposition 4.1.2 and Corollary 4.1.3, we can transfer the problem of existence of optimal solution of the cross-impact LOB model to the problem of existence of optimal strategy in zero-spread LOB model. Proposition 4.1.2 and Corollary 4.1.3 respectively states that for any strategy in the admissible set $\mathcal{A}(Q)$, the zero-spread cost $\mathcal{C}^{\infty}$ is a lower bound of the cross-impact cost $\mathcal{C}^{\beta}$, and furthermore the minimum cost of the zero-spread LOB model is a lower bound of the minimum cost of the cross-impact LOB model.

We closely follow the line offered in Alfonsi and Schied [3] and Alfonsi and Acevdeo [2], where the price impact reversion and volume impact reversion are both considered. Firstly, we use Euler-Lagrange formalism to find the discrete-time optimal trading strategy. Then, by taking each trading interval $[t_i, t_{i+1})$ to be very small approaching zero, one obtains a candidate continuous-time optimal strategy.

Then the verification of optimality is done by direct calculation. Therefore, in the sequel we will discuss the optimal execution strategy under four cases of zero-spread LOB model, namely reversion of the volume impact process in discrete time (E_t Dis) and continuous time (E_t Cts), reversion of the price impact process in discrete (D_t Dis) and continuous time (D_t Cts) setting.

We also obtain sufficient conditions for absence of TTPM under all four cases, which are summarised in Table 4.1. An overview of the contributions in this chapter is: Our results generalise the results in Alfonsi and Acevdeo [2] in terms of optimal strategy (Proposition 4.1.5, 4.1.6, 4.1.10, and 4.1.11) and absence condition of transaction-triggered price manipulation (Corollary 4.1.7 and 4.1.12). For the zero-spread LOB model in Alfonsi and Acevdeo [2], the sufficient conditions on absence of TTPM is independent on shape function. While, in our case with more general shape function, the condition in Corollary 4.1.7 is more restrictive than that in Alfonsi and Acevdeo [2].

Section 4.2 is devoted to constant time varying shape function of cross impact LOB model. There, we will present some numerical examples of our cross-impact LOB optimal execution strategies. It is further assumed that the shape function is of the form $f(t, x) = q(t)$. Comparative analysis on the shape function, same side resilience rate ρ and cross-impact resilience rate β are conducted. Figure 4.4 suggests that the bigger the cross impact resilience rate, the more volatile the optimal strategy alternating between buy and sell. As shown in Figure 4.6, the bigger the same side resilience rate, the more pure buy (sell) orders are used and the smaller size for each child order.

Model:	Conditions for absence of TTPM
E_t Dis	$h_{i+1}(x) = \frac{g(t_i,x) - a_{i+1}g(t_{i+1}, a_{i+1}x)}{1-a_{i+1}}$, for $x \in \mathbb{R}$ and $0 \le i \le N-1$ **Assumption:** For $\forall t \in [0,T]$, the shape function $f(t,x)$ satisfies that: 1) $x\frac{\partial f}{\partial x}(t,x) \le 0$, and 2) $x\eta(t,x) \ge 0$. $a_i + a_{i+1} \le 1$ for $1 \le i \le N-1$
E_t Cts	$h_t(x) = g(t,x) + x\frac{\partial g}{\partial x}(t,x) - \frac{1}{\rho}\frac{\partial g}{\partial t}(t,x)$. **Assumption:** For $\forall t \in [0,T]$, the shape function $f(t,x)$ satisfies that: 1) $x\frac{\partial f}{\partial x}(t,x) \le 0$, and 2) $x\eta(t,x) \ge 0$. $\frac{\partial}{\partial x}\left(\frac{\eta}{f}\right)(t,x) > 0$. $\left[2\rho\zeta_t\frac{\partial f}{\partial t} + 2\frac{\partial f}{\partial t}\eta + \rho f\eta + 2\rho^2\zeta_t f - f\frac{\partial \eta}{\partial t}\right](t, g(t,\zeta_t)) \ge 0$
D_t Dis	$p_{i+1}(x) = x\frac{\frac{1}{a_{i+1}} - a_{i+1}f(t_{i+1},x)\frac{\partial g}{\partial x}\left(t_i, F\left(t_i, \frac{x}{a_{i+1}}\right)\right)}{1 - a_{i+1}f(t_{i+1},x)\frac{\partial g}{\partial x}\left(t_i, F\left(t_i, \frac{x}{a_{i+1}}\right)\right)}$. **Assumption:** For each fixed $t \in [0,T]$ and all $x \in \mathbb{R}$, the shape function $f(t,x)$ needs to satisfy: 1) $x\frac{\partial f}{\partial x}(t,x) \ge 0$; 2) $\frac{\partial \eta}{\partial t}(t,x) \le 0$ and $x\left(f\frac{\partial^2 f}{\partial t\partial x} - \frac{\partial f}{\partial t}\frac{\partial f}{\partial x}\right) \ge 0$; 3) $x \to x\frac{\partial_x f(t,x)}{f(t,x)}$ is non-decreasing on $\mathbb{R}_-$ and non-increasing on $\mathbb{R}_+$.
D_t Cts	$p_t(x) = x\frac{2\rho f(t,x) - \partial_t f(t,x) + \rho x\partial_x f(t,x)}{\rho f(t,x) - \partial_t f(t,x) + \rho x\partial_x f(t,x)}$. **Assumption:** $\rho f(t,x) + \rho x\frac{\partial f}{\partial x}(t,x) - \frac{\partial \eta}{\partial x}(t,x) > 0$, $\rho x\partial_t\left(\frac{\partial_x f}{f}(t,x)\right) - \partial_t\left(\frac{\partial_t f}{f}(t,x)\right) + \left(\rho - \frac{\partial_t f}{f}(t,x)\right)\left(2\rho - \frac{\partial_t f}{f}(t,x)\right) > 0$

Table 4.1: Summary of results under four cases.

4.1 Construction of optimal trading strategy

Let us summarise the optimal execution problem that we consider in this chapter. The cross-impact cost function $\mathcal{C}^\beta(X^A, X^B)$ is defined as

$$\begin{aligned}\mathcal{C}^\beta(X^A, X^B) = \int_0^T (D_t^A - L_t^A) dX_t^A &+ \sum_{t\le T} \Big[G^A(t, E_t^A - J_t^A + \Delta X_t^A) \\ &- G^A(t, E_t^A - J_t^A)\Big] + \int_0^T (D_t^B - L_t^B) dX_t^B + \sum_{t \le T}\Big[G^B(t, E_t^B - J_t^B \\ &- \Delta X_t^B) - G(t, E_t^B - J_t^B)\Big], \end{aligned} \tag{4.1.1}$$

where D^A, D^B are the same side price impact, L_t^A, L_t^B are the cross price impact, E_t^A, E_t^B are the same side volume impact, and J_t^A, J_t^B are the cross volume impact. They are related by formula (3.1.6) in Chapter 3. Recall the following functions too. $g^A(t,x)$ and $g^B(t,x)$ are inverse functions of $F^A(t,x)$ and $F^B(t,x)$ such that $g^A(t, F^A(t,x)) = x$ and $g^B(t, F^B(t,x)) = x$. $F^A(t,x)$ and $F^B(t,x)$ are functions satisfying

$$\begin{aligned}\frac{\partial F^i}{\partial t}(t,x) &= \eta^i(t,x),\\ F^A(t,x) &= \int_0^x f^A(t,y)dy\end{aligned}$$

and

$$F^B(t,x) = \int_{-x}^0 f^B(t,y)dy.$$

where $f^A(t,x)$ is ask side shape function and $f^B(t,x)$ is bid side shape function. The auxiliary function $G^A(t,x)$ and $G^B(t,x)$ are defined as

$$\tilde{F}^A(t,x) = \int_0^x y f^A(t,y)dy, \qquad G^A(t,x) = \tilde{F}^A(t, g^A(t,x)),$$

and

$$\tilde{F}^B(t,x) = \int_{-x}^0 y f^B(t,y)dy, \qquad G^B(t,x) = \tilde{F}^B(t, g^B(t,x)).$$

The admissible set is given by

$$\mathcal{A}(Q) = \Big\{(X^A, X^B) \in \mathbb{R}_+^2 \;\Big|\; X_0^A = X_0^B = 0,\ \int_0^{T+} dX_s^A - \int_0^{T+} dX_s^B = Q,$$
$$X_t^A \geq X_s^A,\ X_t^B \geq X_s^B \text{ for any } t \geq s,$$
$$\text{and } X^A,\ X^B \text{ are bounded variation.}\Big\}. \tag{4.1.2}$$

The admissible set $\mathcal{A}(Q)$ is closed, convex but not bounded. In this case, we make use of the notion of a coercive function.

Definition 4.1.1: *A function $f(x) : \mathbb{R}^n \to \mathbb{R}$ is said to be coercive if for every sequence $\{x^\nu\} \subset \mathbb{R}^n$ for which $\|x^\nu\| \to \infty$ it must be the case that $f(x^\nu) \to +\infty$ as well. That is, for any constant $M > 0$ there is a constant $R(M) > 0$ such that $\|f(x^\nu)\| > M$ whenever $\|x^\nu\| > R(M)$.*

It turns out that Proposition 4.1.2 and Corollary 4.1.3 enable us to simplify the task of proving the existence of optimal solution in cross-impact LOB model.

Proposition 4.1.2: *Given the cost functional $\mathcal{C}^\infty$ for zero-spread LOB model, $\mathcal{C}^\beta$ with $0 < \beta < \infty$ for cross-impact LOB model and $\mathcal{C}^0$ for one-side LOB model, for any strategy $X \in \mathcal{A}(Q)$ these cost functions satisfy the following relationship*

$$\mathcal{C}^\infty(X) \leq \mathcal{C}^\beta(X) \leq \mathcal{C}^0(X).$$

Moreover, for all pure strategies $X \in \mathcal{A}_P(Q)$, the three cost functionals coincide i.e.

$$\mathcal{C}^\infty(X) = \mathcal{C}^\beta(X) = \mathcal{C}^0(X).$$

The optimal execution problem under the zero-spread LOB model is formulated as follows. Minimise the zero-spread cost functional

$$\mathcal{C}^\infty = \int_0^T (D_t^A - D_t^B) dX_t + \sum_{t \in [0,T]} [G(t, E_t^A - E_t^B + \Delta X_t) - G(t, E_t^A - E_t^B)] \tag{4.1.3}$$

over the admissible set

$$\mathcal{A}^{\infty}(Q) := \left\{ X_t : [0, T_+] \to \mathbb{R} \;\middle|\; X_0 = 0, X_{T_+} = Q \text{ and } X_t = \int_0^t dX_s^A - \int_0^t dX_s^B \right.$$
$$\left. \text{for any } (X_t^A, X_t^B) \in \mathcal{A}(Q) \right\}.$$

From the analysis of section 3.2.1, we know the function $G(t,x)$ is defined by $G(t,x) = G^A(t,x)$ and $F(t,x) = \int_0^x f(t,y)dy$, $\eta(t,x) = \frac{\partial F}{\partial t}(t,x)$ and $g(t,x)$ the inverse function of $F(t,x)$ satisfying $g(t, F(t,x)) = x$.

Corollary 4.1.3: *If there is an optimal solution $X^{*,\beta}$ for the optimal execution problem (4.1.1) and an optimal solution $X^{*,\infty}$ for problem (4.1.3), one obtains the following relationship*

$$\mathcal{C}^{\infty}(X^{*,\infty}) \leq \mathcal{C}^{\beta}(X^{*,\beta}) \text{ over } \mathcal{A}(Q).$$

As a result, we can restrict our attention to the existence of optimal solution in zero-spread LOB model. In the discrete time case, we will demonstrate the existence by proving the coerciveness of the cost functionals and then we construct a semi closed-form optimal execution strategy. By taking the time step of trading strategies to zero, we get a candidate solution. In this way, we can transfer the results from discrete to continuous time.

Before stepping into the detailed discussion, we will make a simplification assumption that the same side and opposite resilience rate are constant, i.e. $\rho_t = \rho$ for all $t \in [0,T]$ in this chapter. All results can be extended to the deterministic resilience rate case in the same way. Besides, in the rest of this section only, we will drop the superscription and denote by $\mathcal{C} = \mathcal{C}^{\infty}$.

4.1.1 Model of volume impact reversion in discrete time

Denote by $a_n = e^{-\rho(t_n - t_{n-1})}$ for $1 \leq n \leq N$. The dynamics of the volume impact process in discrete time are summarised by equations

$$E_0 = 0,\ E_n = \sum_{i=0}^{n-1} \xi_i e^{-\rho(t_n - t_i)} \text{ for } n = 1, \dots, N$$

and

$$E_{0+} = \xi_0,\ E_{n+} = \sum_{i=0}^{n-1} \xi_i e^{-\rho(t_n - t_i)} + \xi_n \text{ for } n = 1, \dots, N,$$

where ξ_i is the trade size at each trading time $t_i \in \mathcal{T}$.

We rewrite the zero-spread cost function (4.1.3) in discrete time setting by

$$\begin{aligned}
\mathcal{C}(\xi) &= \sum_{n=0}^{N} [G(t_n, E_n + \xi_n) - G(t_n, E_n)] \\
&= \sum_{n=0}^{N-1} [G(t_n, E_n + \xi_{t_n}) - G(t_{n+1}, E_{n+1})] + G(t_N, E_N + \xi_N) \\
&= \sum_{n=0}^{N-1} [G(t_n, E_n + \xi_n) - G(t_{n+1}, a_{n+1}(E_n + \xi_n))] + G(t_N, E_N + \xi_N). \quad (4.1.4)
\end{aligned}$$

In order to construct a discrete time optimal strategy in Proposition 4.1.5, we need the following Assumption 4.1.4 and auxiliary function:

$$h_{i+1}(x) = \frac{g(t_i, x) - a_{i+1} g(t_{i+1}, a_{i+1} x)}{1 - a_{i+1}}, \text{ for } x \in \mathbb{R} \text{ and } 0 \leq i \leq N-1. \quad (4.1.5)$$

Assumption 4.1.4: *For $\forall t \in [0, T]$, the shape function $f(t, x)$ satisfies the conditions $x \frac{\partial f}{\partial x}(t, x) \leq 0$ and $\eta(t, x) \geq 0$.*

Proposition 4.1.5: *Under Assumption 4.1.4, the auxiliary function (4.1.5) is bijective and we denote by h_i^{-1} its inverse function. Construct a trading strategy ξ^**

as follows

$$\begin{aligned}\xi_0^* &= h_1^{-1}(\nu),\\ \xi_i^* &= h_{i+1}^{-1}(\nu) - a_i h_i^{-1}(\nu), \quad 1 \le i \le N-1 \\ \xi_N^* &= F(t_N,\nu) - a_N h_N^{-1}(\nu),\end{aligned} \tag{4.1.6}$$

where ν *is the unique solution of the following equation*

$$Q = \sum_{i=1}^{N} \xi_{t_i} = (1-a_1)h_1^{-1}(\nu) + \ldots + (1-a_N)h_N^{-1}(\nu) + F(t_N,\nu).$$

Strategy (4.1.6) is the unique minimiser for the cost functional (4.1.4) over $\mathcal{A}_N^\infty(Q)$*. Moreover, this zero-spread LOB model does not admit PMS. The first and the last trades have the same sign as* Q*. The intermediate strategies* ξ_i^**,* $1 \le i \le N-1$ *has the same sign as* Q*, if the following condition*

$$a_i + a_{i+1} \le 1 \textit{ for } 1 \le i \le N-1 \tag{4.1.7}$$

is satisfied.

In other words, Assumption 4.1.4 and condition (4.1.7) are the sufficient conditions for absence of TTPM. The condition (4.1.7) does not depend on the shape function $f(t,x)$, but only on the resilience rate ρ. With this condition, we can tell that the bigger the resilience rate ρ, the less profitable to use the TTPM strategies in this discrete time volume impact reversion model.

The condition (4.1.7) for absence of TTPM strategy is more restrictive than the condition (33) in Theorem 2.3 of Alfonsi and Acevdeo [2], where the shape function takes the form $f(t,x) = q(t)f(x)$. In that case, condition (4.1.7) is satisfied. This is proved in Lemma 4.3.4.

4.1.2 Model of volume impact resilience in continuous time

For continuous time case, we utilise the optimal execution strategy obtained in discrete time set-up, and provide a verification argument that the corresponding continuous analogy strategy is an optimal solution. We propose a continuous time auxiliary function

$$h_t(x) = g(t,x) + x\frac{\partial g}{\partial x}(t,x) - \frac{1}{\rho}\frac{\partial g}{\partial t}(t,x), \tag{4.1.8}$$

We assume the shape function $f(t,x)$ to be such that the auxiliary function (4.1.8) is bijective on $\mathbb{R}$ and has positive first derivative on x is positive, namely $\frac{\partial h_t}{\partial x}(x) > 0$. Later in Corollary 4.1.7 we will show that the auxiliary function $h_t(x)$ is bijective and strictly increasing on $\mathbb{R}$ if Assumption 4.1.4 and condition (4.1.10) are satisfied.

Proposition 4.1.6: *Assume a continuous differentiable shape function $f(t,x)$ such that the auxiliary function (4.1.8) is bijective and $\frac{\partial h_t}{\partial x}(x) > 0$ on $\mathbb{R}$. Denote by $h_t^{-1}(x)$ the inverse function of $h_t(x)$. Construct a trading trading strategy X^* in the following way: Trade at time $t = 0$ a singular order of size ξ_0^*; then continuously trade with rate ξ_t^* on time interval $(0,T)$; finally at time T submit a singular order of ξ_T^*. The trade sizes are determined by*

$$\begin{aligned}\xi_0^* &= \zeta_0,\\ \xi_t^* &= \frac{d\zeta_t}{dt} + \rho\zeta_t\end{aligned}$$

and

$$\xi_T^* = F(T,\nu) - a_N\zeta_T,$$

where we set $\zeta_t := h_t^{-1}(\nu)$ and $\nu \in \mathbb{R}$ uniquely solves the equation

$$Q = \int_0^T \rho h_t^{-1}(\nu)dt + F(T,\nu). \tag{4.1.9}$$

Then the cost function 4.1.3 is non-negative and has a minimiser X^* over $\mathcal{A}^\infty(Q)$. Moreover, the initial trade ξ_0^* and the last trade ξ_T^* have the same sign as Q which *is positive if it is a buy program and is negative if it is a sell program.*

Now let us discuss the sign of the trades given by the optimal strategy. A sufficient condition of absence of TTPM is summarised in Corollary 4.1.7.

Corollary 4.1.7: *Under Assumption 4.1.4, the auxiliary function $h_t(x)$ is continuous differentiable in x, bijective on $\mathbb{R}$ and strictly increasing if*

$$\frac{\partial}{\partial x}\left(\frac{\eta}{f}\right)(t,x) > 0. \tag{4.1.10}$$

Thus, the results of Proposition 4.1.6 hold. Moreover, if

$$\left[2\rho\zeta_t\frac{\partial f}{\partial t} + 2\frac{\partial f}{\partial t}\eta + \rho f\eta + 2\rho^2\zeta_t f - f\frac{\partial \eta}{\partial t}\right](t, g(t,\zeta_t)) \geq 0 \tag{4.1.11}$$

also holds where ζ_t takes values as that in Proposition 4.1.6, ξ_t^* has the same sign as Q for any $0 < t < T$, which excludes TTPM.

We should note that the condition (4.1.11) is not a time continuous analog of the absence condition (4.1.7) in discrete time case and is not a sharp condition as we can see from the proofs. Our results from Corollary 4.1.7 is more restrictive than the condition (22) in Alfonsi and Acevdeo [2] in a way that our results is depend on the shape function. In particular, in Lemma 4.3.6 we provide a proof that our condition (4.1.11) implies the condition (22). Furthermore, it turns out that as long as the shape function takes the form of $f(t,x) = q(t)f(x)$, the condition (22) in Alfonsi and Acevdeo [2] holds under arbitrary function $f(x)$.

4.1.3 Model of price impact reversion in discrete time

Recall that for $1 \leq n \leq N$, we set $a_n = e^{-\rho(t_n - t_{n-1})}$. The dynamics of the price impact process D_t in discrete time are given by

$$D_0 = 0, \qquad D_n = a_n g(t_{n-1}, F(t_{n-1}, D_{n-1}) + \xi_{n-1}) \text{ for } 1 < i \leq N$$

and

$$D_{0+} = g(0, \xi_0), \quad D_{n+} = g(t_n, F(t_n, D_n) + \xi_n) \text{ for } 1 < i \leq N,$$

where $g(t,x)$ is the inverse function of $F(t,x) = \int_0^x f(t,y)dy$ such that $g(t, F(t,x)) = x$.

We translate the cost function (4.1.3) with price impact resilience as

$$\begin{aligned}
\mathcal{C}(\xi) &= \sum_{n=0}^{N} \left[G(t_n, F(t_n, D_n) + \xi_n) - G(t_n, F(t_n, D_n))\right] \\
&= G(t_N, F(t_N, D_N) + \xi_N) + \sum_{n=0}^{N-1} \left[G(t_n, F(t_n, D_n) + \xi_n) - G(t_{n+1}, F(t_{n+1}, D_{n+1}))\right] \\
&= \tilde{F}(t_N, g(t_N, F(t_N, D_N) + \xi_N)) \\
&+ \sum_{n=0}^{N-1} \left[\tilde{F}(t_n, g(t_n, F(t_n, D_n) + \xi_n)) - \tilde{F}(t_{n+1}, a_{n+1} g(t_n, F(t_n, D_n) + \xi_n))\right],
\end{aligned} \tag{4.1.12}$$

where $\tilde{F}$ is defined by $\tilde{F}(t,x) = \int_0^x y f(t,x) dy$. Similarly in order to solve the optimal execution problem in this model, we will need Assumption 4.1.8 and construct an auxiliary function

$$p_{i+1}(x) = x \frac{\frac{1}{a_{i+1}} - a_{i+1} f(t_{i+1}, x) \frac{\partial g}{\partial x}\left(t_i, F\left(t_i, \frac{x}{a_{i+1}}\right)\right)}{1 - a_{i+1} f(t_{i+1}, x) \frac{\partial g}{\partial x}\left(t_i, F\left(t_i, \frac{x}{a_{i+1}}\right)\right)}. \tag{4.1.13}$$

Assumption 4.1.8: *For each fixed $t \in [0,T]$ and all $x \in \mathbb{R}$, the shape function $f(t,x)$ needs to satisfy:*

1. $x \frac{\partial f}{\partial x}(t,x) \geq 0$;
2. $\frac{\partial \eta}{\partial t}(t,x) \leq 0$ *and* $x \left(f \frac{\partial^2 f}{\partial t \partial x} - \frac{\partial f}{\partial t} \frac{\partial f}{\partial x}\right) \geq 0$;
3. $x \to x \frac{\partial_x f(t,x)}{f(t,x)}$ *is non-decreasing on $\mathbb{R}_-$ and non-increasing on $\mathbb{R}_+$.*

Remark 4.1.9: *By analysing the Assumption 4.1.8 item 1 and 3, it is observed that the set of functions satisfying this assumption should be of the form $f(t,x) = \lambda(t)$,* *or equivalently* $\partial_x f(t,x) = 0$ *for all* $x \in \mathbb{R}$. *This is due to the fact that function* $x\frac{\partial_x f(t,x)}{f(t,x)} = 0$ *at* $x = 0$.

This observation is also reflected by the Assumption 2.2 in Alfonsi and Acevdeo [2] and by Corollary 2.6 in Alfonsi and Acevdeo [2]. Assumption 2.2 is a restriction of our Assumption 4.1.8 with the shape function of the form $f(t,x) = f(x)\lambda(t)$. *The functions that satisfy that assumption are constant functions on* $\mathbb{R}$. *The sufficient conditions of absence of TTPM in Corollary 2.6 for general separable shape functions coincides with those for constant shape functions, which is trivial from the Assumption 2.2.*

Nevertheless, when finding optimal solution(s) for discrete time price impact reversion, one should not restrict on the set of constant shape functions. Our proof in the discrete time relies on Lagrange multiplies which requires to show first that the cost function has a minimum. In addition, the candidates of optimal strategy in continuous time are obtained from the discrete time optimal solutions shown in the following Proposition 4.1.10.

When stating the continuous time optimal strategy, we slightly relax Assumption 4.1.8 since the proof of the continuous time case relies on a verification argument. In Example 4.1.13 we will show the existence of optimal solutions of time-varying non-constant shape function LOB models and conditions to exclude the PMS and TTPM.

Proposition 4.1.10: *Under Assumption 4.1.8, the auxiliary function (4.1.13) is bijective and we denote by* p_i^{-1} *its inverse function. Construct a trading strategy* ξ^* *as follows*

$$\begin{aligned}
\xi_0^* &= F\left(t_0, \frac{p_1^{-1}(\nu)}{a_1}\right), \\
\xi_i^* &= F\left(t_i, \frac{p_{i+1}^{-1}(\nu)}{a_{i+1}}\right) - F\left(t_i, p_i^{-1}(\nu)\right), \quad 1 \le i \le N-1 \qquad (4.1.14)
\end{aligned}$$

and

$$\xi_N^* = F(t_N, \nu) - F\left(t_N, p_N^{-1}(\nu)\right),$$

where ν is the unique solution of the following equation

$$Q = \sum_{i=1}^{N} \left[F\left(t_{i-1}, \frac{p_i^{-1}(\nu)}{a_i}\right) - F\left(t_i, p_i^{-1}(\nu)\right) \right] + F(t_N, \nu).$$

Strategy (4.1.14) is the unique minimiser for the cost function (4.1.12) over $\mathcal{A}_N^{\infty}(Q)$. The cost functional $\mathcal{C}(\xi)$ is non-negative and the first and the last trades have the same sign as Q.

Not like in the case of discrete time volume resilience model, it is difficult to find a more restrict condition than $sgn(v)F\left(t_i, \frac{p_{i+1}^{-1}(\nu)}{a_{i+1}}\right) - F\left(t_i, p_i^{-1}(\nu)\right) \geq 0$ to make the trades ξ_i^* have the same sign as Q.

4.1.4 Model of price impact reversion in continuous time

We propose the continuous auxiliary function

$$p_t(x) = x \frac{2\rho f(t,x) - \partial_t f(t,x) + \rho x \partial_x f(t,x)}{\rho f(t,x) - \partial_t f(t,x) + \rho x \partial_x f(t,x)}. \tag{4.1.15}$$

Now we are ready to show that no PMS exists and that there is a unique optimal execution strategy if the auxiliary functions for $t \in [0,T]$ are bijective on $\mathbb{R}$ with a positive first derivative on x.

Proposition 4.1.11: *Assume the shape function $f(t,x) \in C^2([0,T] \times \mathbb{R}, \mathbb{R}^+)$ satisfies the condition*

$$\rho f(t,x) + \rho x \frac{\partial f}{\partial x}(t,x) - \frac{\partial \eta}{\partial x}(t,x) > 0, \tag{4.1.16}$$

and is such that the auxiliary function (4.1.15) is bijective on $\mathbb{R}$ with positive first derivative $\partial_x p_t(x) > 0$. Denote by $p_t^{-1}(x)$ the inverse function of $p_t(x)$. Construct a trading strategy X^* in the following way: Trade at time $t = 0$ a singular order of size ξ_0^*; then continuously trade with rate ξ_t^* in time $(0,T)$; finally at time T submit *a singular order of* ξ_T^*. *The trade sizes are determined by*

$$\xi_0^* = F(t_0, \zeta_0),$$
$$\xi_t^* = f(t, \zeta_t)\left(\frac{d\zeta_t}{dt} + \rho\zeta_t\right)$$

and

$$\xi_T^* = F(t_N, \nu) - F(t_N, \zeta_T),$$

where we set $\zeta_t := p_t^{-1}(\nu)$ *and* $\nu \in \mathbb{R}$ *uniquely solve the equation*

$$Q = F(T, \nu) + \int_0^T \left[\rho p_t^{-1}(\nu) f(t, p_t^{-1}(\nu)) - \eta(t, p_t^{-1}(\nu))\right] dt. \tag{4.1.17}$$

Then the strategy X^* *is the minimiser of the cost function (4.1.3) and PMS does not exist. The initial trade* ξ_0^* *has the same sign as* Q.

In the following Corollary 4.1.12, we show that if the Assumption 4.1.8 holds, the increasing, bijection condition of the auxiliary function p_t is automatically satisfied.

Corollary 4.1.12: *Let* f *be twice continuous differentiable. Under Assumption 4.1.8, if the condition* $\rho f(t,x) + \rho x \frac{\partial f}{\partial x}(t,x) - \frac{\partial \eta}{\partial x}(t,x) > 0$ *holds, the function* p_t *is* C^1, *bijective and strictly increasing. Thus, the results of Proposition 4.1.11 hold and the last trade* ξ_T^* *has the same sign as* Q. *Besides, if*

$$\rho x \partial_t \left(\frac{\partial_x f}{f}(t,x)\right) - \partial_t \left(\frac{\partial_t f}{f}(t,x)\right) + \left(\rho - \frac{\partial_t f}{f}(t,x)\right)\left(2\rho - \frac{\partial_t f}{f}(t,x)\right) > 0 \tag{4.1.18}$$

also holds, ξ_t^* *has the same sign as* Q *for any* $0 < t < T$, *which rules out TTPM.*

As stated in Remark 4.1.9, due to Assumption 4.1.8, the results of Corollary 4.1.12 coincides with the condition (30) in Corollary 2.6 of Alfonsi and Acevdeo [2] for absence of PMS and TTPM. From the proof of this corollary, we should note that the condition (4.1.18) is not sharp. In the following, we study the absence of PMS and existence of optimal solution for a LOB model with a non-separable time-varying shape function.

Example 4.1.13: *Consider a shape function of the form*

$$f(t,x) = tx^2 + a, \quad \textit{for some positive constant } a > 0.$$

In this case, $\frac{\partial f}{\partial t}(t,x) = x^2$ *and* $\frac{\partial f}{\partial x}(t,x) = 2tx$. *The auxiliary function* $p_t(x)$ *and its derivative on* x *are then respectively given by*

$$p_t(x) = x\frac{(4\rho t - 1)x^2 + 2a\rho}{(3\rho t - 1)x^2 + a\rho}$$

and

$$\partial_x p_t(x) = \frac{1}{(3\rho t - 1)x^2 + a\rho}\left[x^4(4\rho t - 1)(3\rho t - 1) + a\rho x^2(6\rho t - 1) + 2a^2\rho^2\right].$$

Thus, $p_t(x)$ *is bijective and strictly increasing if and only if*

$$3\rho t - 1 \geq 0.$$

We can apply the results of Proposition 4.1.11 in this case. The LOB model with this non-separable, time-varying shape function admits a unique optimal execution strategy, which can be numerically computed.

4.2 Numerical examples of discrete optimal trading strategy in cross-impact LOB model

In this section, we are going to look at some numerical examples of optimal trading strategies in discrete time cross-impact LOB model. We will further assume that the shape function takes the form $f^A(t,x) = f^B(t,x) = q(t)$ where the depth function $q(t)$ is assumed to be deterministic, twice continuous differentiable and strictly positive.

Substituting the shape function $q(t)$ into the cross impact cost function (4.1.1), one gets

$$\mathcal{C}^{\beta}(X^A, X^B) = \int_0^T (D_t^A - L_t^A) dX_t^A + \int_0^T (D_t^B - L_t^B) dX_t^B + \sum_{t \le T} \left(\frac{(\Delta X_t^A)^2}{2q(t)} + \frac{(\Delta X_t^B)^2}{2q(t)} \right). \tag{4.2.1}$$

In particular, before specifying the dynamics of price impact and volume impact, we substitute the constant shape function into equation (3.1.6) and obtain the relation equation between the price impact s^A, s^B and volume impact V^A, V^B, which is given by

$$s_t^i = \frac{V_t^i}{q(t)}, \text{ for } i = A, B.$$

By considering a shape function $\hat{q}(t) = q(T-t)$ and trading times $\hat{t}_{N-i} = T - t_i$, It turns out that the price impact reversion model in constant shape model is mathematically the same as the volume impact reversion model. Without loss of generality, we will only derive the optimal problem under the assumption of price impact reversion in the rest of this section.

Let us introduce the following notations:

$$\begin{aligned}
\Theta^A &= (\theta_0^A, \theta_1^A, \dots, \theta_N^A) \in \mathbb{R}^{N+1}, \\
\Theta^B &= (\theta_0^B, \theta_1^B, \dots, \theta_N^B) \in \mathbb{R}^{N+1}, \\
a_{i,j} &:= \frac{e^{-\rho(t_j - t_i)}}{q(t_i)}, \\
\tilde{a}_{i,j} &:= \frac{e^{-(\rho+\beta)(t_j - t_i)}}{q(t_i)}, \\
A &:= a_{i,j} \mathbb{1}_{\{i<j\}} \\
\tilde{A} &:= \tilde{a}_{i,j} \mathbb{1}_{\{i<j\}}, \\
\bar{A} &:= a_{i,j} \mathbb{1}_{\{i<j\}} + \frac{a_{i,j}}{2} \mathbb{1}_{\{i=j\}}, \\
B &= \frac{1}{2}(\bar{A}^T + \bar{A}), \\
D &= A - \tilde{A},
\end{aligned}$$

$$z := (\Theta^A, \Theta^B),$$

and

$$M := \begin{pmatrix} B & -D \\ -D & B \end{pmatrix}.$$

The cost function of a discrete trading strategy (Θ^A, Θ^B) can be expressed as

$$\begin{aligned}
\mathcal{C}^\beta(\Theta^A, \Theta^B) &= \langle \Theta^A, \bar{A}\Theta^A \rangle - \langle \Theta^B, A\Theta^A \rangle + \langle \Theta^B, \tilde{A}\Theta^A \rangle + \langle \Theta^B, \bar{A}\Theta^B \rangle \\
&\quad - \langle \Theta^A, A\Theta^B \rangle + \langle \Theta^A, \tilde{A}\Theta^B \rangle \\
&= \langle \Theta^A, B\Theta^A \rangle - \langle \Theta^B, D\Theta^A \rangle + \langle \Theta^B, B\Theta^B \rangle - \langle \Theta^A, D\Theta^B \rangle \\
&= \langle z, Mz \rangle,
\end{aligned}$$

where $\langle \cdot, \cdot \rangle : \mathbb{R}^{N+1} \times \mathbb{R}^{N+1} \to \mathbb{R}$ is an inner product.

By introducing a skew-symmetric matrix $\bar{M} = \begin{pmatrix} 0 & \frac{D-D^T}{2} \\ \frac{D-D^T}{2} & 0 \end{pmatrix}$, the cost function can be further simplified to be

$$\mathcal{C}^\beta(\Theta^A, \Theta^B) = \langle z, Mz \rangle = \langle z, (M + \bar{M})z \rangle = \frac{1}{2}\langle z, (M + M^T)z \rangle = \frac{1}{2}\langle z, Hz \rangle, \quad (4.2.2)$$

where $H := M + M^T$ is a symmetric matrix. Thus the optimal problem in constant shape cross-impact LOB model now is to minimise the cost functional

$$\langle z, Hz \rangle,$$

subject to $z \geq 0$ and $\langle e,\, z \rangle = Q$, where $e^T = (1, ..., 1, -1, ..., -1)$.

The object of Figure 4.1 is to study the changes of the minimal costs of a cross-impact LOB model against the cross impact resilience rate β. The star line describes the minimal costs under the cross-impact LOB model. The horizontal line is the minimal costs of a zero-spread LOB model with the same parameters, except that it is independent on the cross impact resilience rate. One observes that as β increases from 0 to 50, the cross-impact minimum cost approaches to the zero-spread minimum cost from above. This is agreeable with the definitions of the cost functions in these two models and the Proposition 4.1.2 that $\lim_{\beta\to\infty} \mathcal{C}^{*,\beta} = \mathcal{C}^{*,\infty}$ over $X \in \mathcal{A}(Q)$.

Figure 4.2 are two bar graphs illustrating respectively the optimal strategies under a cross-impact LOB model with volume impact revision and with price impact revision. They are plotted against a regular time grid $t_i = t_0 + i\tau$ with $\tau = \frac{T}{N}$. The bars above the horizontal line $\xi = 0$ stands for buy orders, while the bars under this line for sell orders. The common features of these two plots are: the optimal strategy for both models are not pure strategies; the first and last order are two lumps which are much bigger than the intermediate orders; there are some time in the trading interval when it is better not to trade. Another interesting observation is that the optimal strategies are showing symmetry in the sense that $X_i^s = X_{N-i}^V$ if we denote by X^s and X^V the optimal strategies under price impact reversion and volume impact reversion model respectively.

Next, we will have a look at how different shape functions affect the optimal trading strategy under the cross-impact LOB framework. Keep other parameters constant, four examples of shape functions are tested, e.g. reverting depth $q(t) = 2 + \cos(2\pi t)$, increasing depth $q(t) = 1 + 2t$, constant depth $q(t) = 5$ and decreasing depth $q(t) = \frac{2}{0.5+t}$. Figure 4.3 consist of four plots showing optimal strategies under different shape functions, which are represented by dashed lines. Plot (a), (c) and (d) show lump orders at the beginning and end of trade time. In plot (b) under increasing depth though, only the end order is relatively much bigger than other orders. Focusing on plot (b), (c) and (d), there exists a positive correlation between the order book depth and optimal order sizes. Specifically, when the depth function $q(t)$ is increasing (decreasing, or constant), the intermediate optimal trading sizes are increasing (decreasing, or constant). Compared to the rest of the plots, plot (a)'s optimal strategies show oscillation between buy and sell orders.

Inspired by the observation in Figure 4.3, the existence of the transaction-triggered price manipulations depends on the shape function of the order book.

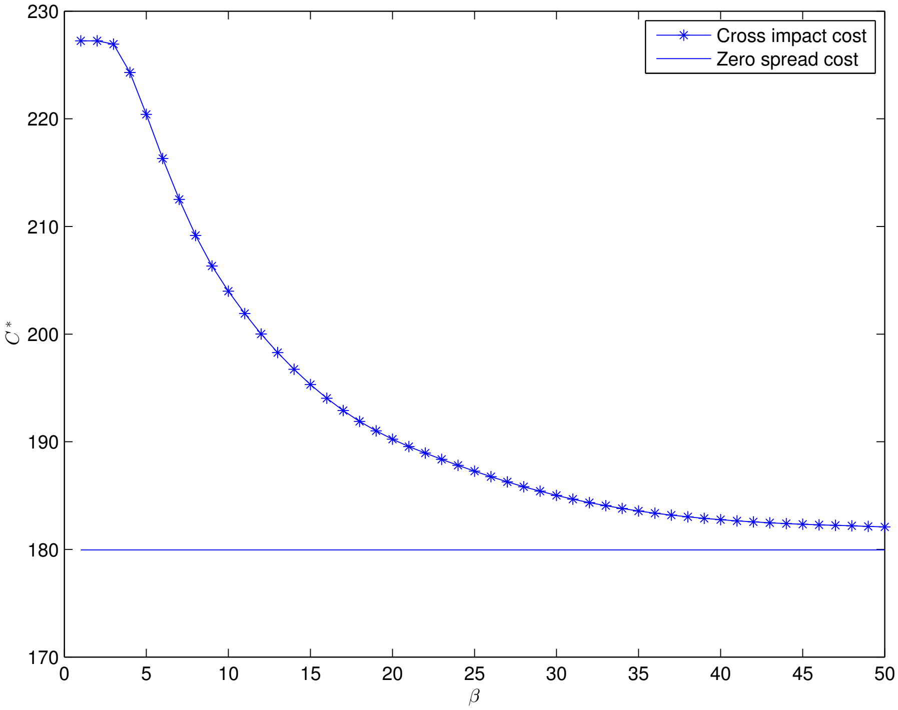
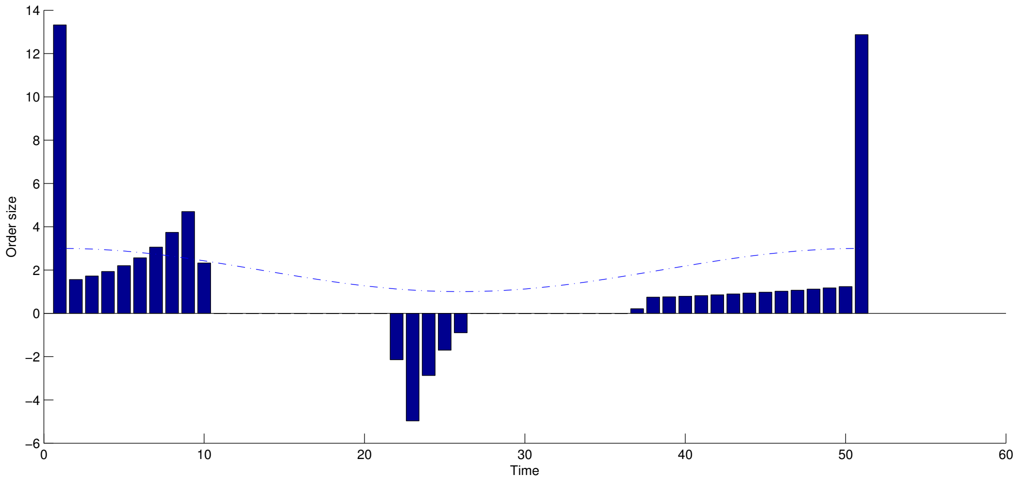
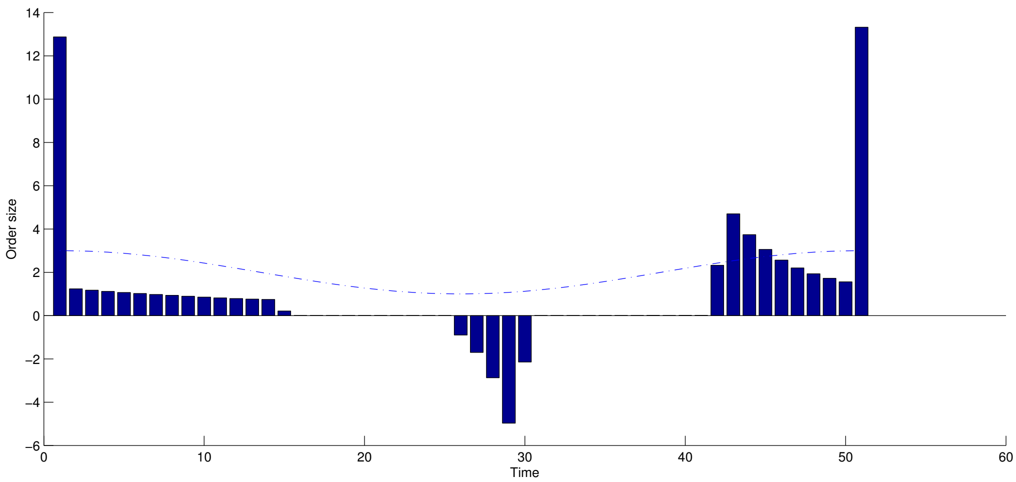


Figure 4.1: The change of the minimal cost $\mathcal{C}^*$ of a cross impact LOB model against the cross impact resilience rate β, with $Q = 50$ shares, $T = 1$, $N = 20$, $\rho = 2$ and $q(t) = 2 + \cos(2\pi t)$. The horizontal line is the minimum cost of a zero-spread LOB model with the same parameters, which is independent on β.

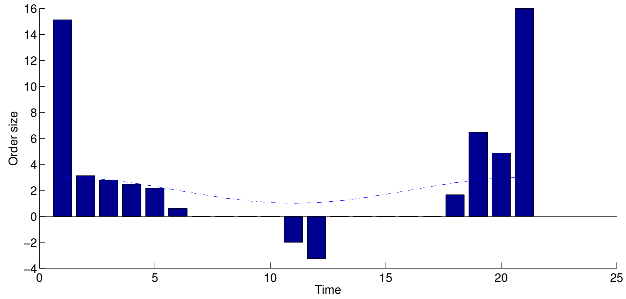


(a) The optimal execution strategies with the volume impact resilience.

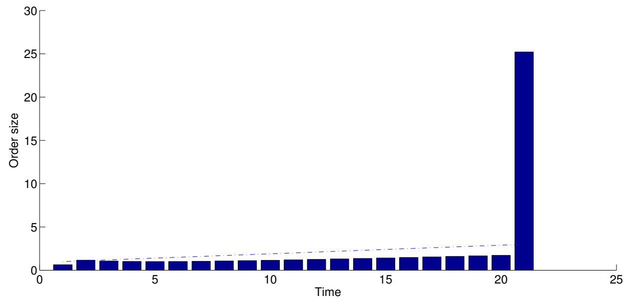


(b) The optimal execution strategies with the price impact resilience.

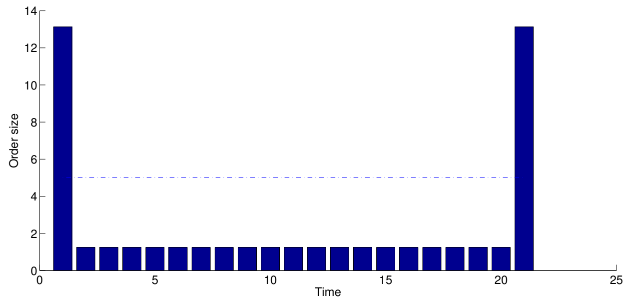
Figure 4.2: The optimal execution strategies for buying 50 shares on a regular time grid with the volume and price impact resilience, with $T = 1$, $N = 50$, $\rho = 2$, $\beta = 10$ and $q(t) = 2 + \cos(2\pi t)$. The dashed line is the plot of the depth function $q(t)$.



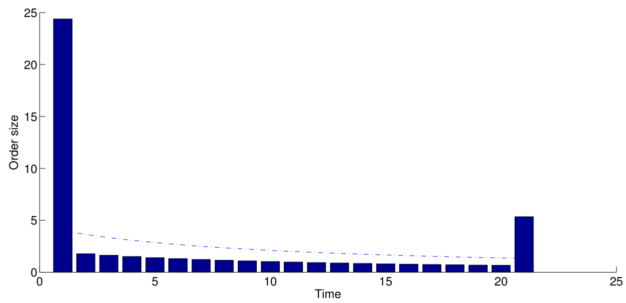
(a) $q(t) = 2 + \cos(2\pi t)$



(b) $q(t) = 1 + 2t$



(c) $q(t) = 5$



(d) $q(t) = \frac{2}{0.5+t}$

Figure 4.3: The optimal execution strategies for buying 50 shares on a regular time grid with four different depth function $q(t)$ (plotted in dashed lines). For each sub-figure, the corresponding depth functions is indicated in its sub-caption. Other parameters are identical for each of the four plots, i.e. $T = 1$, $N = 20$, $\rho = 2$ and $\beta = 5$.

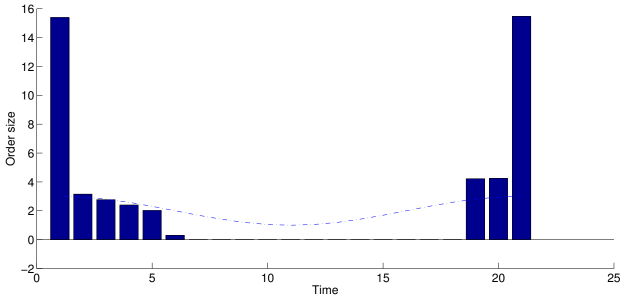
Especially when taking the cross impact resilience rate $\beta = 5$, the optimal strategy is not pure strategy under reverting depth function, meanwhile there are pure buy optimal strategies for increasing, constant and decreasing depth function. Thereafter, in the following Figure 4.4, we fixed the depth function as $q(t) = 2 + \cos(2\pi t)$ and compute the optimal trading orders by setting the cross impact resilience rate β to be $1, 10, 100$ and tends to ∞ respectively. It is shown that when $\beta = 1$, the TTPM can be excluded (by showing pure optimal strategy). As the cross impact resilience rate taking bigger values, more sell orders are used in this purchase program, namely more volatile between buy and sell orders. However, this feature is not true for the model with depth $q(t) = \frac{2}{0.5+t}$ while keeping other parameter the same. In Figure 4.5, there does not exist TTPM strategy for all four values of β.

Finally, we want to study the effect of the same side resilience rate ρ on the optimal strategies. We can achieve this by controlling three parameters at the same time, namely the same side resilience rate ρ, the cross impact resilience rate β and the depth $q(t)$. In Figure 4.6 on the same row of plots, from the left column to the right column, only ρ is increased from 2 to 20. As the same side resilience rate ρ increases, the optimal strategies are less volatile. This is reflected in three aspects: the lump order becomes smaller, as shown in all subplots; the intermediate orders are more evenly distributed, as shown from plot (a) to (b) and from plot (g) to (h); there are less opposite side orders, as shown from plot (c), (e) to plot (d), (f).

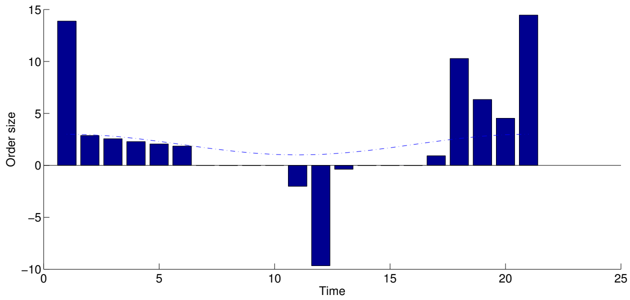
4.3 Proofs

Lemma 4.3.1: *Given the admissible set $\mathcal{A}(Q)$ is defined by equation 4.1.2, $\mathcal{A}(Q)$ is a convex set.*

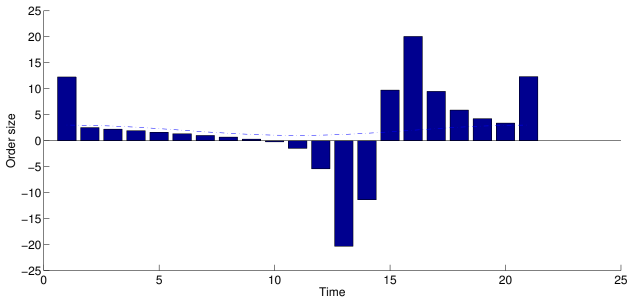
Proof of Lemma 4.3.1. For any (X_1^A, X_1^B), $(X_2^A, X_2^B) \in \mathcal{A}(Q)$ and $\forall \lambda \in (0, 1)$, we check the following three conditions. For $i = A, B$,



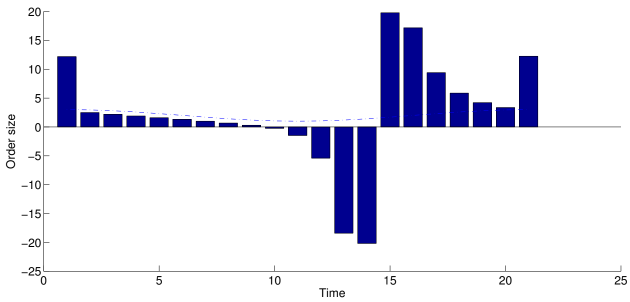
(a) $\beta = 1$



(b) $\beta = 10$

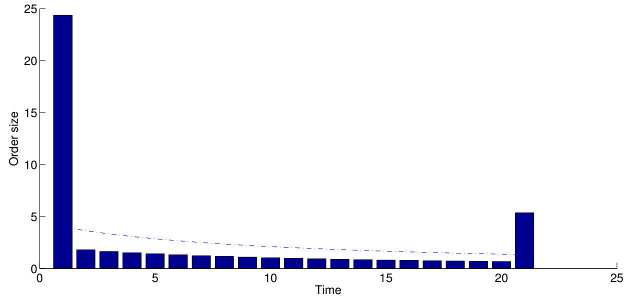


(c) $\beta = 100$

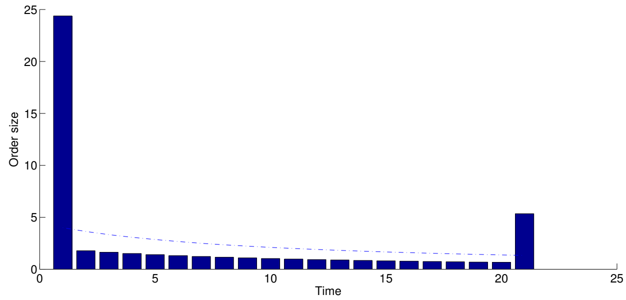


(d) $\beta \to \infty$, i.e. the zero-spread optimal strategy

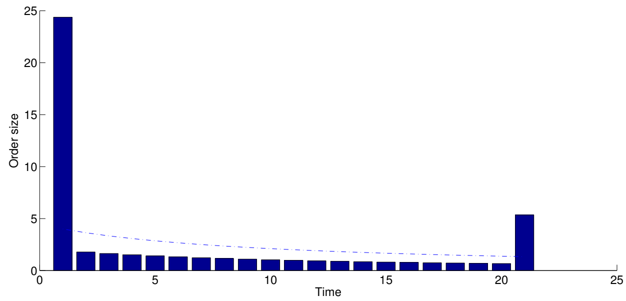
Figure 4.4: The optimal execution strategies for buying 50 shares on a regular time grid with four values of cross impact resilience rate β. For each sub-figure, the corresponding β is indicated in its sub-caption. Other parameters are identical for each of the four plots, i.e. $T = 1$, $N = 20$, $\rho = 2$ and $q(t) = 2 + \cos(2\pi t)$ (plotted in dashed lines).



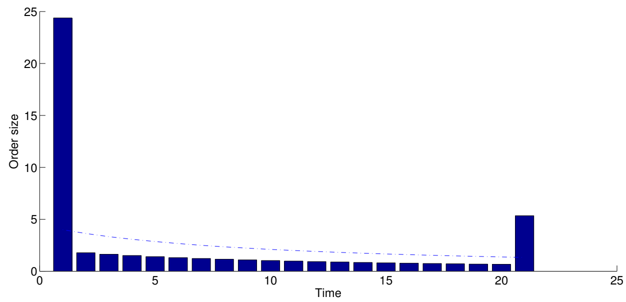
(a) $\beta = 1$



(b) $\beta = 10$

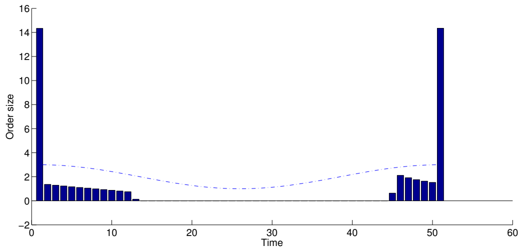


(c) $\beta = 100$

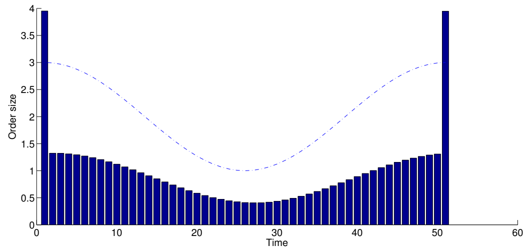


(d) $\beta \to \infty$, i.e. the zero-spread optimal strategy

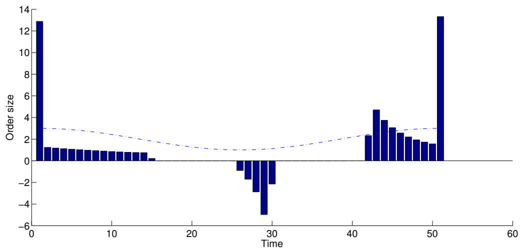
Figure 4.5: The optimal execution strategies for buying 50 shares on a regular time grid with four values of cross impact resilience rate β. For each sub-figure, the corresponding β is indicated in its sub-caption. Other parameters are identical for each of the four plots, i.e. $T = 1$, $N = 20$, $\rho = 2$ and $q(t) = \frac{2}{0.5+t}$ (plotted in dashed lines).



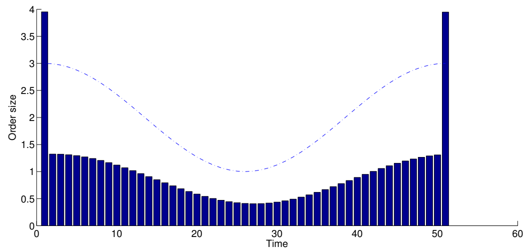
(a) $q(t) = 2 + \cos(2\pi t)$, $\rho = 2$, $\beta = 1$



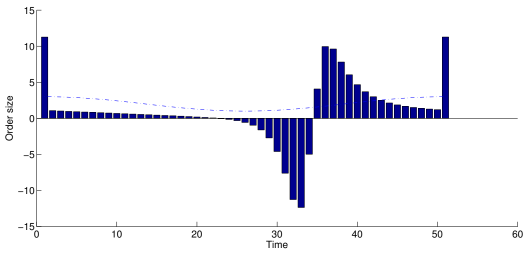
(b) $q(t) = 2 + \cos(2\pi t)$, $\rho = 20$, $\beta = 1$



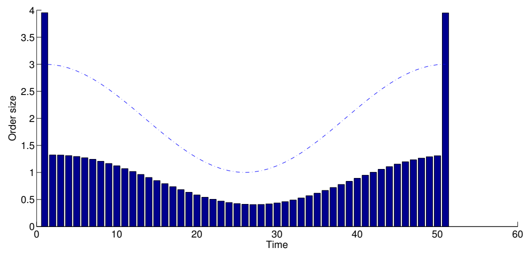
(c) $q(t) = 2 + \cos(2\pi t)$, $\rho = 2$, $\beta = 10$



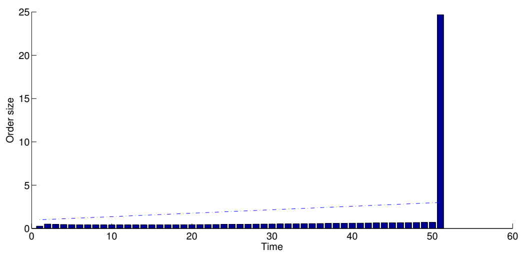
(d) $q(t) = 2 + \cos(2\pi t)$, $\rho = 20$, $\beta = 10$



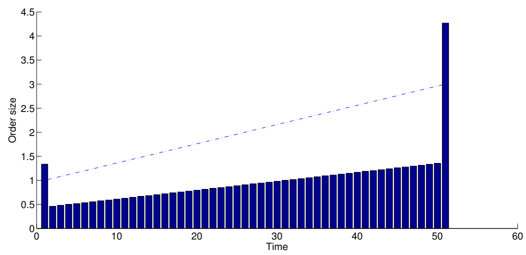
(e) $q(t) = 2 + \cos(2\pi t)$, $\rho = 2$, $\beta \to \infty$



(f) $q(t) = 2 + \cos(2\pi t)$, $\rho = 20$, $\beta \to \infty$



(g) $q(t) = 2t + 1$, $\rho = 2$, $\beta = 10$



(h) $q(t) = 2t + 1$, $\rho = 20$, $\beta = 10$

Figure 4.6: Optimal execution strategy to buy 50 shares on a regular time grid, with $T = 1$, $N = 50$. The dashed lines are the plot of shape functions.

1)

$$\lambda X_1^i(0) + (1-\lambda)X_2^i(0) = 0,$$
$$\lambda X_1(T+) + (1-\lambda)X_2(T+) = \lambda Q + (1-\lambda)Q = Q.$$

2) For any $s < t$, we have $X_1^i(s) \leq X_1^i(t)$ and $X_2^i(s) \leq X_2^i(t)$. Therefore the non-decreasing property holds, i.e.

$$\lambda X_1^i(s) + (1-\lambda)X_2^i(s) \leq \lambda X_1^i(t) + (1-\lambda)X_2^i(t).$$

3) The total variation of $\lambda X_1^i(s) + (1-\lambda)X_2^i(s)$ is also finite and bounded. $\square$

Proof of Proposition 4.1.2. Recall that the cost functions of cross-impact, zero-spread and one-side LOB models are respectively given by

$$\begin{aligned}\mathcal{C}^\beta(X) = &\int_0^T (D_t^A - L_t^A)dX_t^A + \sum_{t\leq T}\left[G^A(t, E_t^A - J_t^A + \Delta X_t^A) - G^A(t, E_t^A - J_t^A)\right] \\ &+ \int_0^T (D_t^B - L_t^B)dX_t^B + \sum_{t\leq T}\left[G^B(t, E_t^B - J_t^B - \Delta X_t^B) - G(t, E_t^B - J_t^B)\right],\end{aligned} \tag{4.3.1}$$

$$\begin{aligned}\mathcal{C}^\infty(X) = &\int_0^T (D_t^A - D_t^B)dX_t^A + \sum_{t\leq T}\left[G^A(t, E_t^A - E_t^B + \Delta X_t^A) - G^A(t, E_t^A - E_t^B)\right] \\ &+ \int_0^T (D_t^B - D_t^A)dX_t^B + \sum_{t\leq T}\left[G^B(t, E_t^B - E_t^A - \Delta X_t^B) - G(t, E_t^B - E_t^A)\right]\end{aligned} \tag{4.3.2}$$

and

$$\begin{aligned}\mathcal{C}^0(X) = &\int_0^T D_t^A dX_t^A + \sum_{t\leq T}\left[G^A(t, E_t^A + \Delta X_t^A) - G^A(t, E_t^A)\right] \\ &+ \int_0^T D_t^B dX_t^B + \sum_{t\leq T}\left[G^B(t, E_t^B - \Delta X_t^B) - G(t, E_t^B)\right].\end{aligned} \tag{4.3.3}$$

For all pure strategies $X \in \mathcal{A}_P(Q)$, the three cost functionals are the same since only either ask side or bid side is involved. The equality part of this proposition is easy to get.

Next, we prove the inequality $\mathcal{C}^{\infty}(X) < \mathcal{C}^{\beta}(X) < \mathcal{C}^{0}(X)$ for $\forall X \in \mathcal{A}(Q) \backslash \mathcal{A}_P(Q)$ in two steps.

Step 1: prove $\mathcal{C}^{\infty} < \mathcal{C}^{\beta}$.

In the case of modelling price impact reversion, given the shape function f^A and f^B, the cross impact resilience rate $0 < \beta < \infty$, the same side resilience rate $0 < \rho_t < \infty$, let us recall the dynamics of same side price impact and cross price impact respectively,

$$\begin{aligned} dD_t^A &= -\rho_t D_t^A dt + \frac{dX_t^A}{f(t, D_t^A)}, \\ dD_t^B &= -\rho_t D_t^B dt + \frac{dX_t^B}{f(t, D_t^B)}, \\ dL_t^A &= -(\beta + \rho_t) L_t^A dt + \beta D_t^B dt, \\ dL_t^B &= -(\beta + \rho_t) L_t^B dt + \beta D_t^A dt. \end{aligned}$$

To simplify the illustration, we introduce the processes $\tilde{D}^A$ and $\tilde{D}^B$ satisfying the equations

$$\begin{aligned} d\tilde{D}_t^A &= -(\rho_t + \beta) \tilde{D}_t^A dt + \frac{dX_t^A}{f(t, D_t^A)}, \\ d\tilde{D}_t^B &= -(\rho_t + \beta) \tilde{D}_t^B dt + \frac{dX_t^B}{f(t, D_t^B)}. \end{aligned}$$

Thereafter, the cross price impact processes L^A and L^B can be expressed by formulas

$$\begin{aligned} L_t^A &= D_t^B - \tilde{D}_t^B, \\ L_t^B &= D_t^A - \tilde{D}_t^A. \end{aligned}$$

Since the admissible strategies X_t^A, X_t^B are non-decreasing and $\tilde{D}^A$, $\tilde{D}^B$ follow exponential decay, one gets the relationships between the cross price impact and same side price impact

$$L_t^A < D_t^B$$

and

$$L_t^B < D_t^A.$$

Thereafter, the ask side price impact $s_t^A = D_t^A - L_t^A$ and the bid side price impact $s_t^B = D_t^B - L_t^B$ satisfy

$$D_t^A - D_t^B < s_t^A \tag{4.3.4a}$$

and

$$D_t^B - D_t^A < s_t^B. \tag{4.3.4b}$$

Via the relationship equation (3.1.6), we can express the ask side volume impact $V_t^A = E_t^A - J_t^A$ and the bid side volume impact $V_t^B = E_t^B - J_t^B$ as

$$V_t^A = F^A(t, s_t^A), \quad V_t^B = F^B(t, s_t^B).$$

Since $F^i(t, x)$ are increasing on x for $i = A, B$, equation (4.3.4) implies that

$$E_t^A - E_t^B = F^A(t, D_t^A - D_t^B) < V_t^A$$

and

$$E_t^B - E_t^A = F^B(t, D_t^B - D_t^A) < V_t^B.$$

In the case of modelling volume impact reversion, given the same side resilience rate ρ_t and cross impact resilience rate $\beta < \infty$, for all $t \in [0, T]$, the dynamics of same side volume impact process and cross volume impact process are given by

$$dE_t^A = -\rho_t E_t^A dt + dX_t^A,$$

$$dE_t^B = -\rho_t E_t^B dt + dX_t^B,$$
$$dJ_t^A = -(\beta + \rho_t) J_t^A dt + \beta E_t^B dt,$$
$$dJ_t^B = -(\beta + \rho_t) J_t^B dt + \beta E_t^A dt.$$

Similarly, we introduce two processes $\tilde{E}^A$ and$\tilde{E}^B$ evolving as follows

$$d\tilde{E}_t^A = -(\rho_t + \beta)\tilde{E}_t^A dt + dX_t^A,$$
$$d\tilde{E}_t^B = -(\rho_t + \beta)\tilde{E}_t^B dt + dX_t^B.$$

Thereafter, the cross price impact processes J^A and J^B can be expressed by formulas

$$J_t^A = E_t^B - \tilde{E}_t^B,$$
$$J_t^B = E_t^A - \tilde{E}_t^A.$$

Since the admissible strategies X_t^A, X_t^B are non-decreasing and $\tilde{E}^A$, $\tilde{E}^B$ follow exponential decay, one gets the relationships between the cross price impact and same side price impact

$$J_t^A < E_t^B$$

and

$$J_t^B < E_t^A.$$

Thereafter, the ask side volume impact $V_t^A = E_t^A - J_t^A$ and the bid side volume impact $V_t^B = E_t^B - J_t^B$ satisfy

$$E_t^A - E_t^B < V_t^A \tag{4.3.5a}$$

and

$$E_t^B - E_t^A < V_t^B. \tag{4.3.5b}$$

Via the relationship equation (3.1.6), we can express the ask side price impact

$s_t^A = D_t^A - L_t^A$ and the bid side price impact $s_t^B = D_t^B - L_t^B$ as

$$V_t^A = g^A(t, s_t^A)$$

and

$$V_t^B = g^B(t, s_t^B).$$

Since $g^i(t,x)$ are increasing on x for $i = A, B$ by Lemma 3.1.4, equation (4.3.5a) implies that

$$D_t^A - D_t^B = g^A(t, E_t^A - E_t^B) < s_t^A$$

and

$$D_t^B - D_t^A = g^B(t, E_t^B - E_t^A) < s_t^B.$$

In both cases, since $G^i(t,x)$ is convex in x by Lemma 3.1.3 and Lemma 3.1.4, one obtains

$$\begin{aligned}&\frac{G^A(t, E_t^A - E_t^B + \Delta X_t^A) - G^A(t, E_t^A - E_t^B)}{\Delta X_t^A}\\ &< \frac{G^A(t, E_t^A - J_t^A + \Delta X_t^A) - G^A(t, E_t^A - J_t^A)}{\Delta X_t^A}\end{aligned}$$

and

$$\begin{aligned}&\frac{G^B(t, E_t^B - E_t^A - \Delta X_t^B) - G^B(t, E_t^B - E_t^A)}{\Delta X_t^B}\\ &< \frac{G^B(t, E_t^B - J_t^B - \Delta X_t^B) - G^B(t, E_t^B - J_t^B)}{\Delta X_t^B}.\end{aligned}$$

From these two inequalities, one obtains $\mathcal{C}^\infty < \mathcal{C}^\beta$.

Step 2: prove $\mathcal{C}^\beta < \mathcal{C}^0$.

Simply since the cross volume impact is positive, i.e. $J_t^i \geq 0$ for $i = A, B$, the convexity of G^i on x implies the following inequalities

$$\frac{G^A(t, E_t^A - J_t^A + \Delta X_t^A) - G^A(t, E_t^A - J_t^A)}{\Delta X_t^A} < \frac{G^A(t, E_t^A + \Delta X_t^A) - G(t, E_t^A)}{\Delta X_t^A}$$

and

$$\frac{G^B(t,E_t^B-J_t^B-\Delta X_t^B)-G^B(t,E_t^B-J_t^B)}{\Delta X_t^B}<\frac{G^B(t,E_t^B-\Delta X_t^B)-G(t,E_t^B)}{\Delta X_t^B}.$$

In addition, since the cross price impact is positive, i.e. $L_t^i \geq 0$ for $i = A, B$, the inequality $\mathcal{C}^\beta < \mathcal{C}^0$ holds. □

Proof of Corollary 4.1.3. Since $X^{*,\infty}$ is a minimiser of $\mathcal{C}^\infty$, one has

$$\mathcal{C}^\infty(X^{*,\infty}) \leq \mathcal{C}^\infty(X^{*,\beta}).$$

At the same time, for strategy $X^{*,\beta}$ Proposition 4.1.2 implies that

$$\mathcal{C}^\infty(X^{*,\beta}) \leq \mathcal{C}^\beta(X^{*,\beta}).$$

Therefore, we obtain $\mathcal{C}^\infty(X^{*,\infty}) \leq \mathcal{C}^\beta(X^{*,\beta})$ over $\mathcal{A}(Q)$. □

Lemma 4.3.2: *Given the volume impact dynamics*

$$E_n=\sum_{i=0}^{n-1}\xi_i e^{-\rho(t_n-t_i)}$$

and the discrete time cost functional (4.1.4), we have $\frac{\partial C}{\partial \xi_N} = g(t_N, E_N + \xi_N)$ *and for* $i = 0, \ldots, N-1$,

$$\frac{\partial \mathcal{C}}{\partial \xi_i}-e^{-\rho(t_{i+1}-t_i)}\frac{\partial \mathcal{C}}{\partial \xi_{i+1}}=g(t_i,E_i+\xi_i)-e^{-\rho(t_{i+1}-t_i)}g(t_{i+1},E_{i+1}). \tag{4.3.6}$$

Proof. We have $\frac{\partial E_n}{\partial \xi_i} = 0$ if $i \geq n$, and $\frac{\partial E_n}{\partial \xi_i} = e^{-\rho(t_n-t_i)}$ if $i < n$. Thus, we get

$$\begin{aligned}\frac{\partial \mathcal{C}}{\partial \xi_i}&=g(t_i,E_i+\xi_i)+\sum_{n=i+1}^{N}e^{-\rho(t_n-t_i)}(g(t_n,E_n+\xi_n)-g(t_n,E_n))\\&=g(t_i,E_i+\xi_i)-e^{-\rho(t_{i+1}-t_i)}g(t_{i+1},E_{i+1})+e^{-\rho(t_{i+1}-t_i)}\Big[g(t_{i+1},E_{i+1}+\xi_{i+1})\\&\quad+\sum_{n=i+2}^{N}e^{-\rho(t_n-t_{i+1})}(g(t_n,E_n+\xi_n)-g(t_n,E_n))\Big]\end{aligned}$$

$$= g(t_i, E_i + \xi_i) - e^{-\rho(t_{i+1}-t_i)} g(t_{i+1}, E_{i+1}) + e^{-\rho(t_{i+1}-t_i)} \frac{\partial \mathcal{C}}{\partial \xi_{i+1}}.$$

$\square$

Lemma 4.3.3: *Under Assumption 4.1.4, one obtains:*

1. *for $i \in \{0, \ldots, N-1\}$, the auxiliary function $h_{i+1}(x)$ by equation 4.1.5 is an increasing bijection on $\mathbb{R}$;*
2. *if condition (4.1.7) holds, we have $sgn(x) h_{i+1}^{-1}(x) \geq sgn(x) a_i h_i^{-1}(x)$ for $i \in \{0, \ldots, N-1\}$;*
3. *$sgn(x) F(t_N, x) \geq sgn(x) a_N h_N^{-1}(x)$.*

Proof of Lemma 4.3.3. We will prove each point of this lemma in sequence.

• point 1:

Since the resilience rate ρ is positive, we have $0 < a_i < 1$. By Assumption 4.1.4, $x\eta(t,x) \geq 0$ implies $f(t_i, x) < f(t_{i+1}, x)$ on $\mathbb{R}$ and $x\frac{\partial f}{\partial x} \leq 0$ implies f is non-decreasing on $\mathbb{R}_-$ and non-increasing on $\mathbb{R}_+$. By Lemma 3.1.4 we know $g(t,x)$ is increasing on x, i.e. $g(t_i, a_{i+1}x) < g(t_i, x)$ on $\mathbb{R}_+$ and $g(t_i, a_{i+1}x) > g(t_i, x)$ on $\mathbb{R}_-$.

Thus we compute $\frac{\partial h_{i+1}}{\partial x}(x)$ and obtain

$$\begin{aligned}
\frac{\partial h_{i+1}}{\partial x}(x) &= \frac{\partial_x g(t_i, x) - a_{i+1}^2 \partial_x g(t_{i+1}, a_{i+1}x)}{1 - a_{i+1}} \\
&= \frac{1}{1 - a_{i+1}} \left[\frac{1}{f(t_i, g(t_i, x))} - a_{i+1}^2 \frac{1}{f(t_{i+1}, g(t_{i+1}, a_{i+1}x))} \right] \\
&\geq \frac{1 - a_{i+1}^2}{1 - a_{i+1}} \frac{1}{f(t_i, g(t_i, x))} > 0.
\end{aligned}$$

• point 2:

We set $\hat{g}(t,x) = \partial_x g(t,x) = 1/f(t, g(t,x))$. By Assumption 4.1.4 and Lemma 3.1.4, we know $\hat{g}$ is positive, non-increasing on $\mathbb{R}_-$ and non-decreasing on $\mathbb{R}_+$. Take $\nu \geq 0$ and $y = h_{i+1}^{-1}(\nu)$. We note that $y \geq 0$ because $h_{i+1}(0) = 0$ and h_{i+1} is increasing by the first point of this lemma. Thus, we have

$$\begin{aligned}
\nu &= \frac{g(t_i, y) - a_{i+1} g(t_{i+1}, a_{i+1} y)}{1 - a_{i+1}} \\
&= g(t_{i+1}, a_{i+1} y) + \frac{g(t_i, y)}{1 - a_{i+1}} \\
&\leq g(t_i, y) + \frac{1}{1 - a_{i+1}} \int_0^y \hat{g}(t_i, r) dr \\
&\leq g(t_i, y) + \frac{1}{1 - a_{i+1}} y \hat{g}(t_i, y) \\
&=: R_{i+1}(y).
\end{aligned}$$

Hence we obtain that R_{i+1} is increasing on $\mathbb{R}_+$ and then $y \geq R_{i+1}^{-1}(\nu)$. Take $z = a_i h_i^{-1}(\nu)$. We have

$$\begin{aligned}
\nu &= \frac{g\left(t_{i-1}, \frac{z}{a_i}\right) - a_i g(t_i, z)}{1 - a_i} \\
&= g(t_i, z) + \frac{g\left(t_{i-1}, \frac{z}{a_i}\right) - g(t_i, z)}{1 - a_i} \\
&\geq g(t_i, z) + \frac{1}{1 - a_i} \int_z^{\frac{z}{a_i}} \hat{g}(t_i, r) dr \\
&\geq g(t_i, z) + \frac{\frac{1}{a_i} - 1}{1 - a_i} z \hat{g}(t_i, z) \\
&=: \bar{R}_i(z).
\end{aligned}$$

Therefore, if

$$\frac{\frac{1}{a_i} - 1}{1 - a_i} \geq \frac{1}{1 - a_{i+1}} \tag{4.3.7}$$

or equivalently the condition (4.1.7) holds, we get that $R_{i+1}(x) \leq \bar{R}_i(x)$ for all $x \geq 0$. That is to say one gets $R_{i+1}^{-1}(x) \geq \bar{R}_i^{-1}(x)$ and therefore

$$y \geq R_{i+1}^{-1}(\nu) \geq \bar{R}_i^{-1}(\nu) \geq z.$$

The same arguments for $\nu \leq 0$ give $y \leq R_{i+1}^{-1}(\nu) \leq \bar{R}_i^{-1}(\nu) \leq z$.

• point 3:

Using the above notations, we have $sgn(x)\bar{R}_N(x) \geq sgn(x)g(t_N, x)$, and therefore we get

$$sgn(\nu)F(t_N, \nu) \geq sgn(\nu)\bar{R}_N^{-1}(\nu) \geq sgn(\nu)z = sgn(\nu)a_N h_N^{-1}(\nu).$$

□

Proof of Proposition 4.1.5. We rewrite the cost function (4.1.4) as follows

$$\begin{aligned}\mathcal{C}(\xi) = G\left(t_N, \sum_{i=0}^{N} \xi_{t_i} e^{-\rho(t_N - t_i)}\right) + \sum_{n=0}^{N-1} \Bigg[G\left(t_n, \sum_{i=0}^{n} \xi_{t_i} e^{-\rho(t_n - t_i)}\right) \\ - G\left(t_{n+1}, e^{-\rho(t_{n+1}-t_n)} \sum_{i=0}^{n} \xi_{t_i} e^{-\rho(t_n - t_i)}\right)\Bigg].\end{aligned}$$

Define the linear map $T_1 : \mathbb{R}^{N+1} \to \mathbb{R}^{N+1}$ by $(T_1\xi)_n = \sum_{i=0}^{n} \xi_{t_i} e^{-\rho(t_n - t_i)}$. Further rewrite the cost function

$$\mathcal{C}(\xi) = G(t_N, (T_1\xi)_N) + \sum_{n=0}^{N-1} [G(t_n, (T_1\xi)_n) - G(t_{n+1}, a_{n+1}(T_1\xi)_n)].$$

Note that T_1 is a linear bijection. Thus one obtains $|T_1\xi| \to +\infty$ as $|\xi| \to +\infty$. Moreover, one can get the following inequality

$$\int_0^{a_{n+1}x} g(t_{n+1}, y)dy \leq a_{n+1} \int_0^{x} g(t_{n+1}, y)dy \leq a_{n+1} \int_0^{x} g(t_n, y)dy.$$

The first $\leq$ sign holds since $G(t,x)$ is convex on x and $\frac{\partial G}{\partial x}(t,x) = g(t,x)$. The second $\leq$ sign holds because of $\eta(t,x) \geq 0$ and $\frac{\partial g}{\partial t}(t,x) = -\frac{\eta(t,x)}{f(t,x)}$ on $\mathbb{R}_+$. Equivalently, that is $G(t_{n+1}, a_{n+1}x) \leq a_{n+1}G(t_n, x)$. We then have $G(t_n, (T_1\xi)_n) - G(t_{n+1}, a_{n+1}(T_1\xi)_n) \geq G(t_n, (T_1\xi)_n)(1 - a_{n+1})$. Since the function G is convex on x and $G(t, 0) = 0$, we have $G(t,x) \geq 0$ for all $x \in \mathbb{R}$ and $\lim\limits_{x \to \infty} G(t,x) = \infty$. Therefore $\mathcal{C}(\xi) \geq 0$ and $\mathcal{C}(\xi) \to +\infty$ as $|\xi| \to +\infty$ since g is increasing and $\lim\limits_{x \to +\infty} F(t,x) = +\infty$.

To find the candidate optimal solution, we consider the Lagrangian L defined as

$$L(X,\nu) = \mathcal{C}(X) - \nu(\sum_{i=0}^{N} \xi_i - Q).$$

By setting $\frac{\partial L}{\partial \xi_i} = 0$, we get the relationships

$$\nu(1-a_{i+1}) = g(t_i, E_{t_i} + \xi_{t_i}) - e^{-\rho(t_{i+1}-t_i)} g(t_{i+1}, E_{t_{i+1}}) \text{ for } i = 0, ..., N \text{ and}$$
$$\nu = g(t_N, E_{t_N} + \xi_{t_N}).$$

Thus the Lagrange multiplier ν satisfies,

$$\nu = h_{i+1}(E_{t_i} + \xi_{t_i}), \text{ for } i = 0, \ldots, N-1, \text{ and } \nu = g(t_N, E_{t_N} + \xi_{t_N}). \tag{4.3.8}$$

By Lemma 4.3.3, we know the auxiliary function $h_i(x)$ is bijective and invertible. Let us denote it by $h_i^{-1}(x)$. Thereafter, we have $E_{t_i} + \xi_{t_i}^* = h_{i+1}^{-1}(\nu)$ and then $E_{t_{i+1}} = a_{i+1} h_{i+1}^{-1}(\nu)$ (the dynamics of volume impact process $E_n = e^{-\rho(t_n - t_{n-1})}(E_{n-1} + \xi_{n-1})$). We then summarise the candidate optimal strategy as

$$\xi_0^* = (h_1)^{-1}(\nu),$$
$$\xi_i^* = (h_{i+1})^{-1}(\nu) - a_i (h_i)^{-1}(\nu), \quad 1 \le i \le N-1$$

and

$$\xi_N^* = F(t_N, \nu) - a_N (h_N)^{-1}(\nu),$$

where ν solves the equation

$$Q = \sum_{i=1}^{N} \xi_{t_i}^* = (1-a_1) h_1^{-1}(\nu) + \ldots + (1-a_N) h_N^{-1}(\nu) + F(t_N, \nu).$$

By Lemma 4.3.3 the right side is an increasing bijection on $\mathbb{R}$, and we deduce that there is only one $\nu \in \mathbb{R}$ which satisfies the above equation. This gives the uniqueness of the minimiser.

Moreover, the functions $g(t,x)$ and $h_i(x)$ vanish at $x = 0$ and ν has the same sign as Q, which implies that ξ_0^* and ξ_N^* have the same sign as Q by Lemma 4.3.3. Besides, by Lemma 4.3.3 if (4.1.7) holds, the trades ξ_i^* have also the same sign as Q. $\square$

Lemma 4.3.4: *Given $x\eta(t,x) \geq 0$ for $x \in \mathbb{R}$, $\forall t \geq 0$ the condition $a_i + a_{i+1} \leq 1$ implies that*

$$\frac{1}{\tilde{a}_i}\frac{1-\tilde{a}_i}{1-a_i} \geq \frac{1-\tilde{a}_{i+1}}{1-a_{i+1}},$$

where $\tilde{a}_i = a_i \frac{q(t_{i-1})}{q(t_i)}$, and $f(t,x) = q(t)f(x)$, $\eta(t) = \frac{q'(t)}{q(t)}$.

Proof. Assume a separable shape function $f(t,x) = q(t)f(x)$. By Assumption 3.2.1 of $\frac{\partial F}{\partial t}(t,x) = \eta(t,x)$, we compute correspondingly $\eta(t,x) = q(t)\eta(t)F(x)$ where $F(x) := \int_0^x f(y)dy$. Since $q(t)$ and $F(x) \geq 0$, we have

$$x\eta(t,x) \geq 0 \quad \Leftrightarrow \quad \eta(t) \geq 0 \quad \Leftrightarrow \quad q(t_i) \geq q(t_{i-1})$$

and then

$$0 < \tilde{a}_i = a_i \frac{q(t_{i-1})}{q(t_i)} \leq a_i < 1.$$

In addition, we note that inequality $a_i + a_{i+1} \leq 1$ is the same as the inequality $\frac{\frac{1}{a_i}-1}{1-a_i} \geq \frac{1}{1-a_{i+1}}$. One then has

$$\frac{\frac{1}{a_i}-1}{1-a_i} \geq \frac{1}{1-a_{i+1}} \quad \Rightarrow \quad \frac{1-a_i}{1-a_{i+1}} \leq \frac{1}{a_i}(1-a_i) \leq \frac{1}{a_i}(1-\tilde{a}_i).$$

Now look at the following inequalities

$$\tilde{a}_i \frac{1-a_i}{1-a_{i+1}} \leq a_i \frac{1-a_i}{1-a_{i+1}} \leq 1-\tilde{a}_i \leq \frac{(1-\tilde{a}_i)}{1-\tilde{a}_{i+1}}.$$

As a result, we get

$$\tilde{a}_i \frac{1-a_i}{1-a_{i+1}} \leq \frac{1-\tilde{a}_i}{1-\tilde{a}_{i+1}} \quad \Leftrightarrow \quad \frac{1}{\tilde{a}_i}\frac{1-\tilde{a}_i}{1-a_i} \geq \frac{1-\tilde{a}_{i+1}}{1-a_{i+1}}.$$

$\square$

Lemma 4.3.5: *Given a time continuous function*

$$h_t(x) = g(t,x) + x\frac{\partial g}{\partial x}(t,x) - \frac{1}{\rho}\frac{\partial g}{\partial t}(t,x)$$

and a time discrete function

$$h_i(x) = \frac{g(t_i,x) - a_{i+1}g(t_{i+1}, a_{i_1}x)}{1 - a_{i+1}},$$

where $t_i \in \{t_0, t_1, ..., t_N\}$ *satisfying* $t_{i+1} - t_i = t_i - t_{i-1}$ *and* $t_N = T$*, and set* $\tau := \frac{T}{N}$*, we have* $\lim_{\tau \to 0} h_i(x) = h_t(x)$*.*

Proof. On time grid $\{t_0, t_1, ..., t_N\}$, we set $a_i \equiv e^{-\rho\tau} := a$. We do the following calculation

$$\begin{aligned}
h_t(x) &= \lim_{\tau\to 0}\frac{g(t,x) - ag(t+\tau,ax)}{1-a} = \lim_{\tau\to 0}\frac{\partial_\tau[g(t,x) - ag(t+\tau,ax)]}{\partial_\tau(1-a)} \\
&= \lim_{\tau\to 0}\frac{\rho e^{-\rho\tau}g(t+\tau,ax) - e^{-\rho\tau}\left[\frac{\partial g}{\partial t}(t+\tau,ax) + \frac{\partial g}{\partial x}(t+\tau,ax)\frac{\partial}{\partial\tau}(e^{-\rho\tau}x)\right]}{\rho e^{-\rho\tau}} \\
&= g(t,x) + x\frac{\partial g}{\partial x}(t,x) - \frac{1}{\rho}\frac{\partial g}{\partial t}(t,x).
\end{aligned}$$

$\square$

Proof of Proposition 4.1.6. Since $h_t(x)$ is bijective, denote its inverse function by h_t^{-1}. For $0 \le t \le T$, set

$$C(t,T,E_t,X_t) = [G(t,\zeta_t) - G(t,E_t)] + \int_t^T g(u,\zeta_u)\xi_u du + [G(T,F(T,\nu)) - G(T,\zeta_T)]$$

where we set $\zeta_u = h_u^{-1}(\nu)$, $\xi_u = \frac{d\zeta_u}{du} + \rho\zeta_u$ and $\nu \in \mathbb{R}$ solves the equation

$$-E_t + \int_t^T \rho h_u^{-1}(\nu)du + F(T,\nu) = Q - X_t. \tag{4.3.9}$$

We should note that the function $\nu \to \int_t^T \rho h_u^{-1}(\nu)du + F(T,\nu)$ is bijective and increasing since $h_t(x)$ and $F(t,x)$ are both increasing and bijective on x. So ν uniquely solves the equation (4.3.9).

Remind that the volume impact function E_t is the solution of $dE_t = dX_t - \rho E_t dt$. Set

$$C_t = \int_0^t g(s, E_s) dX_s + \sum_{0 \le s < t} [G(s, E_s + \Delta X_s) - G(s, E_s)] + C(t, T, E_t, X_t)$$

with $C_T = \mathcal{C}(X)$ and $C_0 = C(0, T, 0, Q)$.

• Step 1: show $C(0,T,0,Q) \geq 0$.

Given the boundary value $C(T,T,E_T,X_T) = G(T, E_T + (Q - X_T)) - G(T, E_T)$, let us substitute ξ_u into $C(t,T,E_t,X_t)$ and then apply integration by parts. We get

$$\begin{aligned}
& C(t,T,E_t,X_t) \\
=& [G(t,\zeta_t) - G(t,E_t)] + [G(T, F(T,\nu)) - G(T,\zeta_T)] \\
& + \int_t^T g(u,\zeta_u) d\zeta_u + \rho \int_t^T g(u,\zeta_u)\zeta_u du \\
=& [G(t,\zeta_t) - G(t,E_t)] + [G(T, F(T,\nu)) - G(T,\zeta_T)] + \rho \int_t^T g(u,\zeta_u)\zeta_u du \\
& + \int_t^T \left[dG(u,\zeta_u) - \frac{\partial G}{\partial u}(u,\zeta_u) du \right] \\
=& - G(t,E_t) + G(T, F(T,\nu)) + \int_t^T \left[\rho g(u,\zeta_u)\zeta_u - \frac{\partial G}{\partial u}(u,\zeta_u) \right] du.
\end{aligned}$$

Since $G(0,E_0) = G(0,0) = 0$ and $G(t,x) \geq 0$, it is sufficient to check that the function $\zeta \to \rho \zeta g(t,\zeta) - \frac{\partial G}{\partial t}(t,\zeta)$ is non-negative.

Since $f(t,x)$ is continuous differentiable, we can do the following calculation

$$\frac{\partial G}{\partial t}(t,x) = \frac{\partial g}{\partial t}(t,x) g(t,x) f(t, g(t,x)) + \int_0^{g(t,x)} y \frac{\partial f}{\partial t}(t,y) dy.$$

By definition we know $g(t,0) = 0$ for all $t \in [0,T]$. Thus the function $\zeta \to \rho\zeta g(t,\zeta) -$

$\frac{\partial G}{\partial t}(t,\zeta)$ vanishes at zero and its derivative is given by

$$\begin{aligned}\frac{\partial}{\partial \zeta}\left(\rho\zeta g(t,\zeta) - \frac{\partial G}{\partial t}(t,\zeta)\right) &= \rho g(t,\zeta) + \rho\zeta\frac{\partial g}{\partial x}(t,\zeta) - \frac{\partial^2 G}{\partial t\partial x}(t,\zeta)\\ &= \rho g(t,\zeta) + \rho\zeta\frac{\partial g}{\partial x}(t,\zeta) - \frac{\partial g}{\partial t}(t,\zeta) = \rho h_t(\zeta),\end{aligned}$$

which has the same sigh as ζ.

• Step 2: show $dC_t \geq 0$, and that $dC_t = 0$ holds only for X^*.

First, we consider the case of a jump $\Delta X_t > 0$. The corresponding change of C_t is given by

$$\Delta C_t = G(t, E_t + \Delta X_t) - G(t, E_t) + C(t+, T, E_{t+}, X_{t+}) - C(t, T, E_t, X_t).$$

Since $E_{t+} - E_t = \Delta X_t$, the solution ν_t of equation (4.3.9) also solves $-E_{t+} + \int_t^T \rho h_u^{-1}(\nu_t)du + F(T,\nu_t) = Q - X_{t+}$ and then $\Delta C_t = 0$. Denote by

$$\begin{aligned}V(t,T,E_t,X_t,\nu) =& - G(t,E_t) + G(T, F(T,\nu))\\ &+ \int_t^T \rho g\left(u, h_u^{-1}(\nu)\right)du - \int_t^T \frac{\partial G}{\partial t}\left(u, h_u^{-1}(\nu)\right)du\end{aligned}$$

and we compute

$$\begin{aligned}&\frac{\partial V}{\partial \nu}(t,T,E_t,X_t,\nu)\\ =& \frac{\partial G}{\partial x}(T,F(T,\nu))\frac{\partial F}{\partial x}(T,\nu) - \int_t^T \frac{\partial^2 G}{\partial t\partial x}\left(u,h_u^{-1}(\nu)\right)\partial_x\left(h_u^{-1}\right)(\nu)du\\ &+ \int_t^T \rho\left[\partial_x\left(h_u^{-1}\right)(\nu) g\left(u,h_u^{-1}(\nu)\right) + h_u^{-1}(\nu)\frac{\partial g}{\partial x}\left(u,h_u^{-1}(\nu)\right)\partial_x(h_u^{-1})(\nu)\right]du\\ =& \rho\nu\partial_x(h_u^{-1})(\nu) + \nu f(T,\nu).\end{aligned}$$

Since $d(E_t - X_t) = \left[\int_t^T \rho\partial_x(h_u^{-1})(\nu_t)du + f(T,\nu_t)\right]d\nu_t - \rho h_t^{-1}(\nu_t)dt$, we could rewrite $\partial_\nu V(t,T,E_t,X_t,\nu) = \rho\nu(h_t^{-1}(\nu)E_t)dt$. The differential of C_t is

$$dC_t = g(t,E_t)dX_t + \frac{dC}{dt}(t,T,E_t,X_t)$$

$$
\begin{aligned}
&= g(t,E_t)dX_t - \frac{\partial G}{\partial t}(t,E_t)dt - \frac{\partial G}{\partial x}(t,E_t)dE_t - \rho g(t,\zeta_t)\zeta_t dt \\
&\quad + \frac{\partial G}{\partial t}(t,\zeta_t)dt + \frac{\partial V}{\partial \nu}(t,T,E_t,X_t,\nu_t)d\nu_t \\
&= -\frac{\partial G}{\partial t}(t,E_t)dt + \rho g(t,E_t)dt - \rho g(t,\zeta_t)\zeta_t dt \\
&\quad + \frac{\partial G}{\partial t}(t,\zeta_t)dt + \rho h_t(\zeta_t)(\zeta_t - E_t)dt \\
&:= \theta_t(\zeta).
\end{aligned}
$$

The function $\theta_t(\zeta) := -\frac{\partial G}{\partial t}(t,E_t) + \rho g(t,E_t)dt - \rho g(t,\zeta)\zeta + \frac{\partial G}{\partial t}(t,\zeta) + \rho h_t(\zeta)(\zeta - E_t)$ vanishes at $\zeta = E_t$ and its derivative $\partial_x \theta_t(\zeta) = -\rho g(t,\zeta) - \rho\zeta\frac{\partial g}{\partial x}(t,\zeta) + \frac{\partial g}{\partial t}(t,\zeta) + \partial_x h_t(\zeta)\rho(\zeta - E_t) + \rho h(\zeta) = \rho\partial_x h_t(\zeta)(\zeta - E_t)$ is positive for $\zeta \neq E_t$. This implies that $\mathcal{C}(X) \geq 0$ for $X \in \mathcal{A}^\infty(Q)$ and $dC_t = 0$ only holds for X^*.

If $\overline{X}$ is another optimal strategy, we necessarily have $\zeta_t = E_t$, $dt - a.e.$. Differentiating $\overline{X}_t - Q - E_t + \int_t^T \rho h_u^{-1}(\nu_t)du + F(T,\nu_t) = 0$ on ν_t, one obtains that $\left[\int_t^T \rho\partial_x h_u^{-1}(\nu_t)du + f(T,\nu_t)\right] d\nu_t = 0$, which implies $d\nu_t = 0$ since $\partial_x h_u^{-1}(x) > 0$ and $f > 0$. Thus we get that $\nu_t = \nu$ where ν is the solution of (4.1.9). Thereafter, we get $\Delta\overline{X}_0 = E_{0+} = (h_0)^{-1}(\nu) = \Delta X_0^*$ and then $\overline{X} = X^*$, which gives the uniqueness of the optimal strategy.

We observe that ν has the same sign as Q and thus $\xi_0^* = \zeta_0 = h_0^{-1}(\nu)$ has the same sign as Q. Remind that $\frac{\partial g}{\partial t} = -\frac{\eta(t,g)}{f(t,g)}$ and $\frac{\partial g}{\partial x} = \frac{1}{f(t,g)}$ which are both positive by assumption. Thus we have $sgn(x)h_t(x) \geq sgn(x)g(t,x)$ which implies that $a_N sgn(x)h_t^{-1}(x) \leq sgn(x)h_t^{-1}(x) \leq sgn(x)F(t,x)$. Thus the last trade $\xi_T^* = F(t,\nu) - a_N h_N^{-1}(\nu)$ has the same sign as Q. □

Proof of Corollary 4.1.7. ζ_t has the same sign as Q, and $g(t,\zeta_t)$ has the same sign as Q since $g(t,0) = 0$ and g is increasing on $\mathbb{R}$. It is sufficient to check that $\xi_t^* \geq 0$.

Since $-x\partial_x f(t,x) \geq 0$ by Assumption 4.1.4 and $\frac{\partial}{\partial x}\left(\frac{\eta}{f}\right) > 0$, dropping the arguments for $g = g(t,x)$ we get

$$
\frac{\partial h_t}{\partial x}(x) = 2\frac{\partial g}{\partial x} - x\frac{1}{(f(t,g))^2}\frac{\partial f}{\partial x}(t,g)\frac{\partial g}{\partial x}
$$

$$
\begin{aligned}
&+\frac{1}{\rho}\frac{\frac{\partial\eta}{\partial x}(t,g)\frac{\partial g}{\partial x}f(t,g)-\eta(t,g)\frac{\partial f}{\partial x}(t,g)\frac{\partial g}{\partial x}}{(f(t,g))^2}\\
&=\left[\frac{2}{f}-\frac{x\partial_x f}{f^3}+\frac{\partial_x\eta f-\eta\partial_x f}{\rho f^3}\right](t,g)\\
&>0.
\end{aligned}
$$

We also have $\frac{d\zeta_t}{dt}=-\frac{1}{\partial_x h_t(\zeta_t)}\frac{dh_t}{dt}(\zeta_t)$. Thus the strategy

$$
\begin{aligned}
\xi_t^* &= \frac{1}{\partial_x h_t(\zeta_t)}\left[-\frac{dh_t}{dt}(\zeta_t)+\rho\zeta_t\partial_x h_t(\zeta_t)\right]\\
&=\frac{1}{\partial_x h_t(\zeta_t)}\left[\frac{\eta(t,g)}{f(t,g)}+\frac{\zeta_t}{f(t,g)^2}\left[\frac{\partial f}{\partial t}(t,g)-\frac{\partial f}{\partial x}(t,g)\frac{\eta(t,g)}{f(t,g)}\right]-\frac{1}{\rho f(t,g)}\left[\partial_t\eta(t,g)\right.\right.\\
&\quad\left.\left.-\partial_t f(t,g)\frac{\eta(t,g)}{f(t,g)}\right]+\frac{\eta(t,g)}{\rho f(t,g)^2}\left[\frac{\partial f}{\partial t}(t,g)-\frac{\partial f}{\partial x}(t,g)\frac{\eta(t,g)}{f(t,g)}\right]+\rho\zeta_t\partial_x h_t(\zeta_t)\right]\\
&=\frac{1}{\partial_x h_t(\zeta_t)}\left[-\frac{\zeta_t\frac{\partial f}{\partial x}(t,g)}{f(t,g)^3}\frac{(\rho\zeta_t+\eta(t,g))^2}{\rho\zeta_t}+\frac{2\frac{\partial f}{\partial t}(t,g)}{f(t,g)^2}\frac{\rho\zeta_t+\eta(t,g)}{\rho}\right.\\
&\quad\left.+\frac{1}{f(t,g)}\frac{\rho\eta(t,g)+2\rho^2\zeta_t-\frac{\partial\eta}{\partial t}(t,g)}{\rho}\right].
\end{aligned}
$$

is non-negative if condition (4.1.11) holds since $\partial_x h_t>0$ and $-\zeta_t\frac{\partial f}{\partial x}(t,g(t,\zeta_t))>0$.

□

Lemma 4.3.6: *Under Assumption 4.1.4, condition (4.1.11) implies*

$$
\frac{d}{dt}\left(\frac{\rho}{2\rho+\eta_t}\right)+\rho\left(\frac{\rho+\eta_t}{2\rho+\eta_t}\right)\geq 0,
$$

where $f(t,x)=q(t)f(x)$, $\eta_t=\frac{q'(t)}{q(t)}$.

Proof. In the case of separable shape function $f(t,x)=q(t)f(x)$, we have

$$
\begin{aligned}
F(t,x)&=q(t)F(x)\text{ with }F(x)=\int_0^x f(y)dy,\\
\eta(t,x)&=q(t)\eta_tF(x),\\
\frac{\partial f}{\partial t}(t,x)&=q(t)\eta_t f(x),\\
\frac{\partial\eta}{\partial t}(t,x)&=\left(\eta_t'+\eta_t^2\right)q(t)F(x)
\end{aligned}
$$

and

$$g(t,x) = F^{-1}\left(\frac{x}{q(t)}\right).$$

Recall also the notation in Alfonsi and Acevdeo [2] that $\bar{\zeta}_t := h_{V,t}^{-1}(\nu)$ where ν is the unique solution of $\int_0^T q(t)\rho h_{V,t}^{-1}(\nu)dt + q(T)F(\nu) = Q$ and the auxiliary function is given by $h_{V,t}^{-1}(x) = F^{-1}(x) + \frac{\rho+\eta_t}{\rho}\frac{x}{f(F^{-1}(x))}$.

Next, we will show $h_t(q(t)x) = h_{V,t}(x)$, or equivalently $\zeta_t = q(t)\bar{\zeta}_t$. Substituting $\frac{x}{q(t)} = y$ into $h_t(x)$, we get

$$\begin{aligned}
h_t(q(t)y) &= g(t,x) + x\frac{1}{f(t,g(t,x))} + \frac{1}{\rho}\frac{\eta(t,g(t,x))}{f(t,g(t,x))} \\
&= F^{-1}(y) + q(t)y\frac{1}{q(t)f(F^{-1}(y))} + \frac{1}{\rho}\frac{q(t)\eta_t F(F^{-1}(y))}{q(t)f(F^{-1}(y))} \\
&= h_{V,t}(y).
\end{aligned}$$

Now we can conclude that

$$\begin{aligned}
&\left[2\rho\zeta_t\frac{\partial f}{\partial t} + 2\frac{\partial f}{\partial t}\eta + \rho f\eta + 2\rho^2\zeta_t f - f\frac{\partial \eta}{\partial t}\right](t, g(t,\zeta_t)) \\
&= 2\rho q^2(t)\bar{\zeta}_t\eta_t f(\bar{\zeta}_t) + 2q^2(t)\eta_t^2\bar{\zeta}_t f(\bar{\zeta}) + 2\rho^2 q(t)\bar{\zeta}_t f(\bar{\zeta}_t) \\
&\quad + \rho q^2(t)\eta_t\bar{\zeta}_t f(\bar{\zeta}_t) - q^2(t)\bar{\zeta}_t(\eta_t' + \eta_t^2)f(\bar{\zeta}_t) \\
&= \bar{\zeta}_t f(\bar{\zeta}_t)q^2(t)(2\rho\eta_t + \eta_t^2 + 2\rho^2 + \rho\eta_t - \eta_t') \\
&= \bar{\zeta}_t f(\bar{\zeta}_t)q^2(t)\left[(\rho+\eta_t)(2\rho+\eta_t) - \eta_t'\right] \geq 0 \\
&\Leftrightarrow \frac{d}{dt}\left(\frac{\rho}{2\rho+\eta_t}\right) + \rho\left(\frac{\rho+\eta_t}{2\rho+\eta_t}\right) \geq 0.
\end{aligned}$$

□

Lemma 4.3.7: *Given the cost function (4.1.12), for $i = 0, ..., N-1$, we have*

$$\begin{aligned}
\frac{\partial \mathcal{C}}{\partial \xi_i} &= g(t_i, F(t_i, D_i) + \xi_i) \\
&\quad + \left[\frac{\partial \mathcal{C}}{\partial \xi_{i+1}} - D_{i+1}\right] f(t_{i+1}, D_{i+1})a_{i+1}\frac{\partial g}{\partial x}(t_i, F(t_i, D_i) + \xi_i).
\end{aligned}$$

Proof. The derivative $\frac{\partial D_n}{\partial \xi_i}$ is given by

$$\frac{\partial D_n}{\partial \xi_i} = \begin{cases} 0, & \text{if } i \geq n \\ a_n \frac{\partial g}{\partial x}(t_{n-1}, F(t_{n-1}, D_{n-1}) + \xi_{n-1}), & \text{if } i = n-1 \end{cases}.$$

If $1 \leq i \leq n-2$, we can calculate that

$$\begin{aligned} \frac{\partial D_n}{\partial \xi_i} &= a_n \frac{\partial g}{\partial x}(t_{n-1}, F(t_{n-1}, D_{n-1}) + \xi_{n-1}) f(t_{n-1}, D_{n-1}) \frac{\partial D_{n-1}}{\partial \xi_i} \\ &= a_{i+1} f(t_{i+1}, D_{i+1}) \frac{\partial g}{\partial x}(t_i, F(t_i, D_i) + \xi_i) \frac{\partial D_n}{\partial \xi_{i+1}}. \end{aligned}$$

Thereafter, we have

$$\begin{aligned} \frac{\partial \mathcal{C}}{\partial \xi_i} &= g(t_i, F(t_i, D_i) + \xi_i) + \sum_{n=i+1}^{N} \left[g(t_n, F(t_n, D_n) + \xi_n) - D_n \right] f(t_n, D_n) \frac{\partial D_n}{\partial \xi_i} \\ &= g(t_i, F(t_i, D_i) + \xi_i) \\ &\quad + \left[g(t_{i+1}, F(t_{i+1}, D_{i+1}) + \xi_{i+1}) - D_{i+1} \right] f(t_{i+1}, D_{i+1}) a_{i+1} \frac{\partial g}{\partial x}(t_i, F(t_i, D_i) + \xi_i) \\ &\quad + \left[\frac{\partial \mathcal{C}}{\partial \xi_{i+1}} - g(t_{i+1}, F(t_{i+1}, D_{i+1}) + \xi_{i+1}) \right] a_{i+1} f(t_{i+1}, D_{i+1}) \frac{\partial g}{\partial x}(t_i, F(t_i, D_i) + \xi_i) \\ &= g(t_i, F(t_i, D_i) + \xi_i) + \left[\frac{\partial \mathcal{C}}{\partial \xi_{i+1}} - D_{i+1} \right] f(t_{i+1}, D_{i+1}) a_{i+1} \frac{\partial g}{\partial x}(t_i, F(t_i, D_i) + \xi_i). \end{aligned}$$

□

Lemma 4.3.8: *Under Assumption 4.1.8, we have that:*

1. *the function $x \to x f(t, x)$ is increasing on $\mathbb{R}$, or equivalently $\tilde{F}(t, x) = \int_0^x y f(t, y) dy$ is convex on x;*

2. *$f\left(t_i, \frac{x}{a_{i+1}}\right) - a_{i+1} f(t_{i+1}, x) > 0$ for $i = 0, \ldots, N-1$;*

3. *The auxiliary function*

$$p_{i+1}(x) = x \frac{a_{i+1}^{-1} - a_{i+1} \frac{f(t_{i+1}, x)}{f\left(t_i, \frac{x}{a_{i+1}}\right)}}{1 - a_{i+1} \frac{f(t_{i+1}, x)}{f\left(t_i, \frac{x}{a_{i+1}}\right)}}$$

is well-defined, bijective and increasing and satisfies $sgn(x)p_i(x) \geq |x|$.

Proof. • point 1: Since $f(t,x) > 0$ for $\forall x \in \mathbb{R}$ and $\forall t \in [0,T]$ and $x\partial_x f(t,x) \geq 0$ by Assumption 4.1.8, one gets

$$\frac{\partial}{\partial x} x f(t,x) = f(t,x) + x\partial_x f(t,x) > 0.$$

• point 2: Since $\frac{\partial \eta}{\partial x}(t,x) \leq 0$ by Assumption 4.1.8, we have

$$f\left(t_i, \frac{x}{a_{i+1}}\right) \geq f\left(t_{i+1}, \frac{x}{a_{i+1}}\right).$$

Also, the assumption that $x\frac{\partial f}{\partial x}(t,x) \geq 0$ and $a_{i+1} < 1$ for $i = 0, ..., N-1$ implies

$$f\left(t_{i+1}, \frac{x}{a_{i+1}}\right) \geq f(t_{i+1}, x) > a_{i+1} f(t_{i+1}, x).$$

As a result, we get

$$f\left(t_i, \frac{x}{a_{i+1}}\right) > a_{i+1} f(t_{i+1}, x).$$

• point 3: Denote by $H := \frac{f(t_{i+1},x)}{f\left(t_i, \frac{x}{a_{i+1}}\right)}$ and we know $a_{i+1}H \leq 1$ by the second point of this lemma. Rewrite the auxiliary function (4.1.13) and we get

$$p_{i+1}(x) = x\left[1 + \frac{a_{i+1}^{-1} - 1}{1 - a_{i+1}H}\right].$$

Since $a_{i+1} < 1$, it is easy to check that p_i is well-defined and satisfies $sgn(x)p_i(x) \geq |x|$. Then we compute the derivative of $p_i(x)$ as

$$\begin{aligned}
\partial_x p_i(x) &= \left(1 + \frac{a_{i+1}^{-1} - 1}{1 - a_{i+1}H}\right) + x\frac{1 - a_{i+1}}{a_{i+1}} \frac{a_{i+1}}{(1 - a_{i+1}H)^2}\partial_x H \\
&= \frac{a_{i+1}}{(1 - a_{i+1}H)^2}\left[(1 - a_{i+1}H)(1 - a_{i+1}^2 H) + x(1 - a_{i+1})a_{i+1}\partial_x H\right].
\end{aligned}$$

The derivative $\partial_x H$ is given by

$$\partial_x H = \partial_x \left(\frac{f(t_{i+1}, x)}{f\left(t_i, \frac{x}{a_{i+1}}\right)} \right) = \frac{\partial_x f(t_{i+1}, x) f\left(t_i, \frac{x}{a_{i+1}}\right) - \frac{1}{a_{i+1}} f(t_{i+1}, x) \partial_x f\left(t_i, \frac{x}{a_{i+1}}\right)}{f\left(t_i, \frac{x}{a_{i+1}}\right)^2}.$$

Thereafter, it is sufficient to check that $x\partial_x H \geq 0$, or equivalently

$$x \frac{\partial_x f(t_{i+1}, x)}{f(t_{i+1}, x)} \geq \frac{x}{a_{i+1}} \frac{\partial_x f\left(t_i, \frac{x}{a_{i+1}}\right)}{f\left(t_i, \frac{x}{a_{i+1}}\right)}.$$

Since $a_{i+1} < 1$, one has

$$x \frac{\partial_x f(t_{i+1}, x)}{f(t_{i+1}, x)} \geq \frac{x}{a_{i+1}} \frac{\partial_x f(t_{i+1}, x/a_{i+1})}{f(t_{i+1}, x/a_{i+1})}.$$

Moreover, since $x\left(\frac{\partial^2 f}{\partial x \partial t} - \frac{\partial f}{\partial t}\frac{\partial f}{\partial x}\right) \geq 0$ by assumption, we have $x\partial_t \left(\frac{\partial_x f}{f}\right) \geq 0$. Thus, one could derive that

$$\frac{x}{a_{i+1}} \frac{\partial_x f\left(t_{i+1}, \frac{x}{a_{i+1}}\right)}{f\left(t_{i+1}, \frac{x}{a_{i+1}}\right)} \geq \frac{x}{a_{i+1}} \frac{\partial_x f\left(t_i, \frac{x}{a_{i+1}}\right)}{f\left(t_i, \frac{x}{a_{i+1}}\right)}.$$

□

Proof of Proposition 4.1.10. Since $\tilde{F}(t, x)$ is convex on x and $\tilde{F}(t, 0) = 0$, we have

$\tilde{F}(t_{n+1}, a_{n+1} g(t_n, F(t_n, D_n) + \xi_n)) \leq a_{n+1} \tilde{F}(t_{n+1}, g(t_n, F(t_n, D_n) + \xi_n))$ for all $x \in \mathbb{R}$

Moreover, we have $\tilde{F}(t_n, x) \geq \tilde{F}(t_{n+1}, x)$ by the second condition 2) in Assumption 4.1.8. Thus, the cost function (4.1.12) satisfies

$$\mathcal{C}(\xi) \geq \tilde{F}(t_n, g(t_N, F(t_N, D_N) + \xi_N)) + \sum_{n=0}^{N-1} \tilde{F}(t_n, g(t_n, F(t_n, D_n)) + \xi_n)(1 - a_{n+1}).$$

Set $T_2(\xi) = (\xi_0, F(t_1, D_1) + \xi_1, \ldots, F(t_N, D_N) + \xi_N)$. Since $\lim_{x\to\infty} F(t, x) = \infty$, one has that $|T_2(\xi)| \to \infty$ as $|\xi| \to \infty$, which immediately implies that $\mathcal{C}(\xi) \to \infty$ as

$|\xi| \to \infty$.

We denote by ν a Lagrange multiplier such that

$$\nu = p_{i+1}(D_{i+1}), \text{ for } i = 0, \dots, N-1 \text{ and } \nu = g(t_N, F(t_N, D_N) + \xi_N).$$

By Lemma 4.3.8, we know that p_i is invertible and we denote by p_i^{-1} its inverse function. We then get

$$\begin{aligned}
\xi_0^* &= F\left(t_0, \frac{p_1^{-1}(\nu)}{a_1}\right), \\
\xi_i^* &= F\left(t_i, \frac{p_{i+1}^{-1}(\nu)}{a_{i+1}}\right) - F\left(t_i, p_i^{-1}(\nu)\right), \text{ for } i = 1, \dots, N-1
\end{aligned}$$

and

$$\xi_N^* = F(t_N, \nu) - F\left(t_N, p_N^{-1}(\nu)\right).$$

Besides, ν solves the equation

$$F(t_N, \nu) + \sum_{i=1}^{N} \left[F\left(t_{i-1}, \frac{p_i^{-1}(\nu)}{a_i}\right) - F\left(t_i, p_i^{-1}(\nu)\right) \right] = Q. \tag{4.3.10}$$

Since F and p_i is increasing and bijective in x and the function $y \to F\left(t_{i-1}, \frac{y}{a_i}\right) - F(t_i, y)$ is increasing (its derivative is positive by Lemma 4.3.8), ν is the unique solution to equation (4.3.10), and has the same sign as Q. Thus ξ^* is the unique optimal solution. Moreover, the initial and last trade have the same sign as Q since $sgn(x)p_i(x) \geq |x|$. □

Lemma 4.3.9: *Given a time continuous function*

$$p_t(x) = x \frac{2\rho f(t,x) - \partial_t f(t,x) + \rho x \partial_x f(t,x)}{\rho f(t,x) - \partial_t f(t,x) + \rho x \partial_x f(t,x)}$$

and a time discrete function

$$p_{i+1}(x) = x \frac{\frac{1}{a_{i+1}} - a_{i+1} f(t_{i+1}, x) \frac{\partial g}{\partial x}\left(t_i, F\left(t_i, \frac{x}{a_{i+1}}\right)\right)}{1 - a_{i+1} f(t_{i+1}, x) \frac{\partial g}{\partial x}\left(t_i, F\left(t_i, \frac{x}{a_{i+1}}\right)\right)},$$

where $t_i \in \{t_0, t_1, ..., t_N\}$ *satisfying* $t_{i+1} - t_i = t_i - t_{i-1}$ *and* $t_N = T$*, and set* $\tau := \frac{T}{N}$*, we have* $\lim_{\tau \to} p_i(x) = p_t(x)$*.*

Proof. In a regular time grid $\{t_0, t_1, ..., t_N\}$ such that $t_{i+1}-t_i = t_i-t_{i-1}$ and $t_N = T$, set $\tau := \frac{T}{N}$ and $a \equiv e^{-\rho\tau}$. We do the following calculation

$$
\begin{aligned}
p_t(x) &= x \lim_{\tau\to 0} \frac{a^{-1} - af(t+\tau, x)\frac{\partial g}{\partial x}\left(t, F\left(t, \frac{x}{a}\right)\right)}{1 - af(t+\tau, x)\frac{\partial g}{\partial x}\left(t, F\left(t, \frac{x}{a}\right)\right)} \\
&= x \lim_{\tau\to 0} \frac{a^{-1} - a\frac{f(t+\tau,x)}{f\left(t,\frac{x}{a}\right)}}{1 - a\frac{f(t+\tau,x)}{f\left(t,\frac{x}{a}\right)}} \\
&= x \lim_{\tau\to 0} \frac{f(t, e^{\rho\tau}x)e^{\rho\tau} - f(t+\tau, x)e^{-\rho\tau}}{f(t, e^{\rho\tau}x) - f(t+\tau, x)e^{-\rho\tau}} \\
&= x \lim_{\tau\to 0} \frac{\rho e^{\rho\tau} f(t, e^{\rho\tau}x) + e^{\rho\tau}\partial_x f(t, e^{\rho\tau}x)\rho e^{\rho\tau}x + \rho e^{-\rho\tau} f(t+\tau, x) - \partial_t f(t+\tau, x)e^{-\rho\tau}}{\partial_x f(t, e^{\rho\tau}x)\rho e^{\rho\tau}x + \rho e^{-\rho\tau} f(t+\tau, x) - \partial_t f(t+\tau, x)e^{-\rho\tau}} \\
&= x \frac{2\rho f(t,x) - \partial_t f(t,x) + \rho x \partial_x f(t,x)}{\rho f(t,x) - \partial_t f(t,x) + \rho x \partial_x f(t,x)}.
\end{aligned}
$$

□

Proof of Proposition 4.1.11. Since $p_t(s)$ is bijective, we can denote by $p^{-1}(x)$ its inverse function. For $0 \le t \le T$, we introduce

$$
C(t, T, D_t, X_t) = [\tilde{F}(t, \zeta_t) - \tilde{F}(t, D_t)] + \int_t^T \zeta_u \xi_u du + [\tilde{F}(T, \nu) - \tilde{F}(T, \zeta_T)], \tag{4.3.11}
$$

where $\zeta_u = p_u^{-1}(\nu), \xi_u = f(u, \zeta_u)\left(\frac{d\zeta_u}{du} + \rho\zeta_u\right)$ and $\nu \in \mathbb{R}$, solving the equation

$$
-E_t + \int_t^T \left[\rho p_u^{-1}(\nu) f\left(u, p_u^{-1}(\nu)\right) - \eta\left(u, p_u^{-1}(\nu)\right)\right] du + F(T, \nu) = Q - X_t. \tag{4.3.12}
$$

We should note that the function $x \to \rho x f(t,x) - \eta(t,x)$ is increasing and bijective, since its derivative is equal to $\rho f(t,x) + \rho x\frac{\partial f}{\partial x}(t,x) - \frac{\partial \eta}{\partial x}(t,x)$ and is positive by assumption. Thus ν uniquely solves equation (4.3.12).

Recall that the price impact function D_t and volume impact function E_t

satisfy $E_t = F(t, D_t)$ and $dD_t = -\rho D_t dt$. So we can solve E_t satisfying

$$\begin{aligned}
dE_t &= dX_t + \frac{\partial F}{\partial t}(t, D_t)dt + \frac{\partial F}{\partial x}(t, D_t)dD_t \\
&= dX_t + [\eta(t, D_t) - \rho D_t f(t, D_t)]\, dt \\
&= dX_t + \eta(t, g(t, E_t))dt - \rho g(t, E_t) f(t, g(t, E_t))dt.
\end{aligned}$$

We set

$$C_t = \int_0^t D_s dX_s + \sum_{0 \le s < t} [G(s, E_s + \Delta X_s) - G(s, E_s)] + C(t, T, D_t, X_t)$$

with $C_T = \mathcal{C}(X)$ and $C_0 = C(0, T, 0, Q)$.

• Step 1: show $C(0, T, 0, Q) \ge 0$.

Given that $C(T, T, D_T, X_T) = G(T, Q - X_T + E_T) - G(T, E_T)$, let us substitute ξ_u into $C(t, T, D_t, X_t)$ and then apply integration by parts. We get

$$C(t, T, D_t, X_t) = \tilde{F}(T, \nu) - \tilde{F}(t, D_t) + \int_t^T \left[\rho \zeta_u^2 f(u, \zeta_u) - \frac{\partial \tilde{F}}{\partial t}(u, \zeta_u)\right] du. \quad (4.3.13)$$

Since $\tilde{F}(0, D_0) = \tilde{F}(0, 0) = 0$ and $\tilde{F}(t, x) \ge 0$, it is sufficient to show that the function $\zeta \to \rho f(t, \zeta)\zeta^2 - \frac{\partial \tilde{F}}{\partial t}(t, \zeta)$ is non-negative.

It is obvious that this function vanishes at $\zeta = 0$. Its derivative is equal to $2\zeta\rho f(t, \zeta) + \rho\zeta^2 \frac{\partial f}{\partial x}(t, \zeta) - \zeta \frac{\partial \eta}{\partial x}(t, \zeta)$ and has the same sign as ζ by assumption.

• Step 2: show $dC_t \ge 0$, and that $dC_t = 0$ holds only for X^*.

We first consider the case of a jump $\Delta X_t > 0$. The corresponding change of C_t is given by

$$\Delta C_t = [G(t, E_t + \Delta X_t) - G(t, E_t)] + C(t, T, D_{t+}, X_{t+}) - C(t, T, D_t, X_t).$$

Since $E_{t+} - E_t = \Delta X_t$, the solution ν_t of equation (4.3.12) also solves

$$-E_{t+} + \int_t^T \left[\rho p_u^{-1}(\nu_t) f\left(u, p_u^{-1}(\nu_t)\right) - \eta\left(u, p_u^{-1}(\nu_t)\right)\right] du + F(T, \nu_t) = Q - X_{t+}.$$

Thus $\Delta C_t = 0$ since $\tilde{F}(t, D_t) = G(t, E_t)$. Denote by

$$\begin{aligned} V(t, T, D_t, X_t, v) &= \tilde{F}(T, v) - \tilde{F}(t, D_t) \\ &\quad + \int_t^T \left[\rho\left(p_u^{-1}(v)\right)^2 f\left(u, p_u^{-1}(v)\right) - \frac{\partial \tilde{F}}{\partial t}\left(u, p_u^{-1}(v)\right)\right] du \end{aligned}$$

and we compute

$$\begin{aligned} &\frac{\partial V}{\partial v}(t, T, D_t, X_t, v) \\ &= \partial_x \tilde{F}(T, v) + \int_t^T \partial_x\left(p_u^{-1}\right)(v)\left[2\rho p_u^{-1}(v) f\left(u, p_u^{-1}(v)\right)\right] du \\ &\quad + \int_t^T \partial_x\left(p_u^{-1}\right)(v)\left[\rho\left(p_u^{-1}(v)\right)^2 \partial_x f\left(u, p_u^{-1}(v)\right) - \partial_x \partial_t \tilde{F}\left(u, p_u^{-1}(v)\right)\right] du \\ &= v f(T, v) + \int_t^T \partial_x\left(p_u^{-1}\right)(v) p_u^{-1}(v)\left[2\rho f\left(u, p_u^{-1}(v)\right)\right] du \\ &\quad + \int_t^T \partial_x\left(p_u^{-1}\right)(v) p_u^{-1}(v)\left[\rho p_u^{-1}(v) \partial_x f\left(u, p_u^{-1}(v)\right) - \partial_t f\left(u, p_u^{-1}(v)\right)\right] du \\ &= v\left[f(T, v) + \int_t^T \partial_x\left(p_u^{-1}\right)(v)\left(\rho f\left(u, p_u^{-1}(v)\right)\right) du\right] \\ &\quad + v\left[\int_t^T \partial_x\left(p_u^{-1}\right)(v)\left(\rho p_u^{-1}(v) \partial_x f\left(u, p_u^{-1}(v)\right) - \partial_t f\left(u, p_u^{-1}(v)\right)\right) du\right]. \end{aligned}$$

Since $d(E_t - X_t) = [\eta(t, D_t) - \rho D_t f(t, D_t)]dt$, we get from equation (4.3.12) that

$$\begin{aligned} &\left[\int_t^T \partial_x\left(p_u^{-1}\right)(v)\left[\rho f\left(u, p_u^{-1}(v)\right) + \rho p_u^{-1}(v) \partial_x f\left(u, p_u^{-1}(v)\right) - \partial_x \eta\left(u, p_u^{-1}(v)\right)\right] du\right] dv_t \\ &+ \left[\int_t^T \partial_x F(T, v) du\right] dv_t - \left[\rho p_t^{-1}(v) f\left(t, p_t^{-1}(v)\right) - \eta\left(t, p_t^{-1}(v)\right)\right] dt \\ &= [\eta(t, D_t) - \rho D_t f(t, D_t)]dt. \end{aligned} \tag{4.3.14}$$

Thereafter, we can rewrite $\partial_v V(t,T,D_t,X_t,v)$ as

$$\frac{\partial \tilde{C}}{\partial v}(t,T,D_t,X_t,v)dv_t = v_t[\eta(t,D_t) - \eta(t,\zeta_t) + \rho\zeta_t f(t,\zeta_t) - \rho D_t f(t,D_t)].$$

Since $dD_t = -\rho_t D_t dt + \frac{dX_t}{f(t,D_t)}$, the differential of C_t is given by

$$\begin{aligned}
dC_t &= D_t dX_t - \frac{\partial \tilde{F}}{\partial t}(t,D_t)dt + \rho D_t^2 f(t,D_t)dt - D_t dX_t - \rho\zeta_t^2 f(t,\zeta_t)dt \\
&\quad + \frac{\partial \tilde{F}}{\partial t}(t,\zeta_t)dt + \frac{\partial \tilde{C}}{\partial v}(t,T,D_t,X_t,v)dv_t \\
&= \frac{\partial \tilde{F}}{\partial t}(t,\zeta_t)dt - \frac{\partial \tilde{F}}{\partial t}(t,D_t)dt + \rho D_t^2 f(t,D_t)dt - \rho\zeta_t^2 f(t,\zeta_t)dt \\
&\quad + p_t(\zeta_t)[\eta(t,D_t) - \eta(t,\zeta_t) + \rho\zeta_t f(t,\zeta_t) - \rho D_t f(t,D_t)]dt \\
&=: \psi_t(\zeta_t)dt.
\end{aligned}$$

We have $\psi_t(D_t) = 0$ and can compute its derivative as

$$\partial_x\psi_t(\zeta) = \partial_x p_t(\zeta)[\eta(t,D_t) - \eta(t,\zeta) + \rho\zeta f(t,\zeta) - \rho D_t f(t,D_t)].$$

Since $\partial_x p_t(\zeta) > 0$, it is sufficient to look at the term $\alpha_t(x) := \eta(t,D_t) - \eta(t,x) + \rho x f(t,x) - \rho D_t f(t,D_t)$. We then compute its derivative $\partial_x\alpha_t(x)$ given by $\rho x\partial_x f(t,x) + \rho f(t,x) - \partial_x\eta(t,x)$. The function $\alpha_t(x)$ is positive on $\zeta > D_t$, and negative on $\zeta < D_t$ since $\partial_x\alpha_t(x)$ is positive by assumption. Thus, D_t is the unique minimiser of ψ_t: $\psi_t(D_t) = 0$ and $\psi_t(\zeta) > 0$ for $\zeta \neq D_t$.

If X is an optimal strategy, we necessarily have $\zeta_t = D_t\ dt - a.e.$. From equation (4.3.14), one obtains that

$$\begin{aligned}
\Bigg[\int_t^T \partial_x\left(p_u^{-1}\right)(\nu)\big[\rho f\left(u,p_u^{-1}(\nu)\right) + \rho_u p_u^{-1}(\nu)\partial_x f\left(u,p_u^{-1}(\nu)\right) \\
- \partial_x\eta\left(u,p_u^{-1}(\nu)\right)\big]du\Bigg]d\nu_t = 0,
\end{aligned}$$

which implies $d\nu_t = 0$ since $\partial_x p_u^{-1}(x) > 0$ and $\partial_x\alpha_t(x) > 0$. Thus we get that $\nu_t = \nu$ where ν is the solution of (4.3.12). Thereafter, we get $\Delta X_0 = F(t_0,\zeta_0) = \Delta X_0^*$ and then $X = X^*$, which gives the uniqueness of the optimal strategy.

Since ν has the same sign as Q, $p_t(x)$ is bijective and $F(t,x)$ is increasing along $\mathbb{R}$, one obtains that $\xi_0^* = F(t_0, \zeta_0)$ has the same sign as Q. $\square$

Proof of Corollary 4.1.12. Set $\Delta := \rho\left(1 + \frac{x\partial_x f(t,x)}{f(t,x)}\right) - \frac{\partial_t f(t,x)}{f(t,x)}$. First, dropping the arguments for $f(t,x)$, we can rewrite the auxiliary function $p_t(x)$ as

$$\begin{aligned} p_t(x) &= x\frac{2\rho f - \partial_t f + \rho x \partial_x f}{\rho f - \partial_t f + \rho x \partial_x f} \\ &= x\left(\frac{\rho}{\rho + x\frac{\partial_x f}{f} - \frac{\partial_t f}{f}} + 1\right) = x\left(\frac{\rho}{\Delta} + 1\right). \end{aligned}$$

We then compute its derivative

$$\begin{aligned} \frac{\partial p_t}{\partial x}(x) &= \left(1 + \frac{\rho}{\Delta}\right) - \frac{x\rho}{\Delta^2}\frac{\partial \Delta}{\partial x} = \frac{1}{\Delta^2}(\Delta^2 + \rho\Delta - x\rho\frac{\partial \Delta}{\partial x}) \\ &= \frac{1}{\Delta^2}\left[\Delta^2 + \rho\Delta - \rho^2 x \partial_x\left(\frac{x\partial_x f}{f}\right) + \rho x \partial_x\left(\frac{\partial_t f}{f}\right)\right]. \end{aligned}$$

By Assumption 4.1.16, we have $x\partial_x\left(\frac{x\partial_x f}{f}\right) \leq 0$, $x\frac{\partial}{\partial x}\left(\frac{\partial_t f}{f}\right) \geq 0$ and also we assume $\Delta > 0$, which implies $\partial_x p_t(x) > 0$. Since $\Delta > 0$ and $p_t(0) = 0$, we also have p_t is bijective and $sgn(x)p_t(x) \geq |x|$. We then could deduce that $sgn(x)p_t^{-1}(x) \leq |x|$, which implies that the last trade ξ_T^* has the same sign as Q since $F(t,x)$ is increasing in x.

We also know $\frac{d\zeta_t}{dt} = -\frac{1}{\partial_x p_t(\zeta_t)}\frac{dp_t}{dt}(\zeta_t)$ and the first derivative

$$\begin{aligned} \frac{d}{dt}p_t(x) &= -\frac{x\rho}{\Delta^2}\frac{\partial \Delta}{\partial t} \\ &= -\frac{x\rho}{\Delta^2}\left[\rho x \partial_t\left(\frac{\partial_x f}{f}\right) - \partial_t\left(\frac{\partial_t f}{f}\right)\right]. \end{aligned}$$

Thus the strategy

$$\begin{aligned} \xi_t^* &= f\left(t, \zeta_t\right)\left(\frac{d\zeta_t}{dt} + \rho\zeta_t\right) \\ &= f(t,\zeta_t)\left(\frac{1}{\partial_x p_t(\zeta_t)}\frac{\zeta_t\rho}{\Delta^2}\left[\rho\zeta_t\partial_t\left(\frac{\partial_x f}{f}\right) - \partial_t\left(\frac{\partial_t f}{f}\right)\right] + \rho\zeta_t\right) \end{aligned}$$

$$
\begin{aligned}
&= \frac{\rho\zeta_t f(t,\zeta_t)}{\partial_x p_t(\zeta_t)\Delta^2}\left[\rho\zeta_t\partial_t\left(\frac{\partial_x f}{f}\right) - \partial_t\left(\frac{\partial_t f}{f}\right) + \partial_x p_t(\zeta_t)\Delta^2\right] \\
&= \frac{\rho\zeta_t f(t,\zeta_t)}{\partial_x p_t(\zeta_t)\Delta^2}\left[\rho\zeta_t\partial_t\left(\frac{\partial_x f}{f}\right) - \partial_t\left(\frac{\partial_t f}{f}\right) + \Delta^2 + \rho\Delta - \rho^2\zeta_t\partial_x\left(\frac{x\partial_x f}{f}\right) + \rho\zeta_t\partial_x\left(\frac{\partial_t f}{f}\right)\right] \\
&\geq \frac{\rho\zeta_t f(t,\zeta_t)}{\partial_x p_t(\zeta_t)\Delta^2}\left[\rho\zeta_t\partial_t\left(\frac{\partial_x f}{f}\right) - \partial_t\left(\frac{\partial_t f}{f}\right) + \Delta^2 + \rho\Delta\right] \\
&\geq \frac{\rho\zeta_t f(t,\zeta_t)}{\partial_x p_t(\zeta_t)\Delta^2}\left[\rho\zeta_t\partial_t\left(\frac{\partial_x f}{f}\right) - \partial_t\left(\frac{\partial_t f}{f}\right) + \left(\rho - \frac{\partial_t f}{f}\right)\left(2\rho - \frac{\partial_t f}{f}\right)\right] \\
&\geq 0.
\end{aligned}
$$

The first inequality holds by the last two items of Assumption 4.1.8 and the second inequality does by the first item of Assumption 4.1.8. Moreover, the last inequality holds if condition (4.1.18) holds. □

Lemma 4.3.10: *Under Assumption 4.1.8, if* $\rho\left(1 + \frac{x\partial_x f(t,x)}{f(t,x)}\right) - \frac{\partial_t f(t,x)}{f(t,x)} \geq 0$ *condition (4.1.18) implies*

$$
\frac{d}{dt}\left(\frac{\rho - \eta_t}{2\rho - \eta_t}\right) + \rho\left(\frac{\rho - \eta_t}{2\rho - \eta_t}\right) \geq 0
$$

where $f(t,x) = q(t)f(x)$ *and* $\eta_t = \frac{q'(t)}{q(t)}$.

Proof. In the case of separable shape function $f(t,x) = q(t)f(x)$, we have

$$
\begin{aligned}
F(t,x) &= q(t)F(x) \text{ with } F(x) = \int_0^x f(y)dy, \\
\eta(t,x) &= q(t)\eta_t F(x), \\
\frac{\partial f}{\partial t}(t,x) &= q(t)\eta_t f(x), \\
\frac{\partial \eta}{\partial t}(t,x) &= \left(\eta_t' + \eta_t^2\right) q(t)F(x)
\end{aligned}
$$

and

$$
g(t,x) = F^{-1}\left(\frac{x}{q(t)}\right).
$$

Now we can conclude that

$$
\rho x\partial_t\left(\frac{\partial_x f}{f}\right) - \partial_t\left(\frac{\partial_t f}{f}\right) + \left(\rho - \frac{\partial_t f}{f}\right)\left(2\rho - \frac{\partial_t f}{f}\right)
$$

$$
\begin{aligned}
&= \; -\eta_t' + (\rho - \eta_t)(2\rho - \eta_t) \geq 0 \\
&\Leftrightarrow \frac{d}{dt}\left(\frac{\rho - \eta_t}{2\rho - \eta_t}\right) + \rho\left(\frac{\rho - \eta_t}{2\rho - \eta_t}\right) \geq 0.
\end{aligned}
$$

□

Chapter 5

Application II: Absence and presence of market irregularity conditions

Another application of our cross impact LOB model is to study the market impact model regularities when one cannot get closed-form optimal solutions of the execution problem. Our goal is to understand whether and how the dynamics of the LOB, specifically the depth and resilience of the order book, will create market irregularities. In particularly, we will research how the cross impact resilience factor β affects the market irregularity absence and presence conditions.

Some theoretical studies of market manipulations in market impact models, such as Huberman and Stanzl [48], Alfonsi et al. [6], Gatheral [32] and Klöck [52], address the market irregularity problem merely on models without spread. Fruth et al. [31] argues that in models without spread it might appear that there are price manipulations while in practice the spread precludes these price manipulations. At first sight, this statement is quite trivial since traders would not make profits from a round-trip if there is a cost to cross the spread. However, as we can see in the sequel, our cross impact model possesses non-zero spread too but it is possible that all three market irregularities exist.

All of these inspire us to think of the inclusion of two sides resilience into the analysis of market irregularity issues and think of the profitability depending on two sides resilience of the order book. Indeed, the cross impact LOB model provides us new insights about the market irregularity. Our Proposition 5.1.1, Proposition 5.1.2, Proposition 5.2.1, Proposition 5.3.1 and Proposition 5.3.2 show that the market irregularity conditions in models without spread is weaker than that in the cross-impact LOB model and the conditions in cross-impact model is weaker than that in one-side LOB model. All the results about this weaker relationship between zero-spread and cross-impact are due to the argument we made in Proposition 4.1.2 that for an arbitrary strategy $X \in \mathcal{A}(Q)$, the costs under zero-spread model is always less than or equal to the costs under the cross-impact model.

Moreover, we provide necessary conditions for absence of PMS, TTPM and existence of PLC under the cross-impact LOB model. Particularly, the necessary condition (5.3.2) becomes sufficient for zero-spread models.

In this chapter, we assume a constant time-varying shape function of the form $f^A(t,x) = f^B(t,x) = q(t)$ unless specifically stated. The depth function $q(t) : [0,T] \to (0,\infty)$ is deterministic and twice continuous differentiable. For simplicity, the same side resilience rate ρ and cross impact resilience rate β are assumed to be constant. In this constant shape function setting-up, as discussed in section 4.2, we will focus on the models with price impact resilience. We remind the reader the main notations here.

Given a constant shape function $q(t)$, the same side price impact D_t^A, D_t^B and cross price impact L_t^A, L_t^B take the following forms:

$$
\begin{aligned}
D_t^A(X_t^A) &= \int_{[0,t)} \frac{e^{-\rho(t-s)}}{q(s)} dX_s^A, \\
D_t^B(X_t^B) &= \int_{[0,t)} \frac{e^{-\rho(t-s)}}{q(s)} dX_s^B,
\end{aligned}
$$

$$L_t^B(X_t^A) = \int_{[0,t)} \frac{e^{-\rho(t-s)}}{q(s)} \left(1 - e^{-\beta(t-s)}\right) dX_s^A$$

and

$$L_t^A(X_t^B) = \int_{[0,t)} \frac{e^{-\rho(t-s)}}{q(s)} \left(1 - e^{-\beta(t-s)}\right) dX_s^B.$$

The cost function of the cross impact LOB model is

$$\mathcal{C}^\beta = \int_0^T (D_t^A - L_t^A) dX_t^A + \int_0^T (D_t^B - L_t^B) dX_t^B + \sum_{t \le T} \left(\frac{(\Delta X_t^A)^2}{2q(t)} + \frac{(\Delta X_t^B)^2}{2q(t)} \right). \tag{5.0.1}$$

The cost function of the one-side LOB model is

$$\mathcal{C}^0 = \int_0^T D_t^A dX_t^A + \int_0^T D_t^B dX_t^B + \sum_{t \le T} \left(\frac{(\Delta X_t^A)^2}{2q(t)} + \frac{(\Delta X_t^B)^2}{2q(t)} \right). \tag{5.0.2}$$

The cost function of the zero-spread LOB model is

$$\mathcal{C}^\infty = \int_0^T D_t dX_t + \sum_{t \le T} \frac{\Delta X_t^2}{2q(t)}, \tag{5.0.3}$$

where

$$D_t := \int_{[0,t)} \frac{e^{-\rho(t-s)}}{q(s)} dX_s^A - \int_{[0,t)} \frac{e^{-\rho(t-s)}}{q(s)} dX_s^B$$

and

$$X_t := \int_{[0,t)} dX_s^A - \int_{[0,t)} dX_s^B.$$

The structure of this chapter is: section 5.1, 5.2 and 5.3 present conditions for the absence and presence of three irregularities TTPM, PMS and PLC respectively. We will review the definition of each market irregularity first, and then look at how each of them is affected by our cross-impact resilience rate β. It is followed by absence and presence conditions of each irregularity. The section 5.4 show some examples of price impact dynamics generated by three different LOB models which admits PMS.

5.1 Absence of transaction-triggered price manipulation strategies (TTPM)

Recall that a market impact model does not admit transaction-triggered price manipulation (TTPM) if for any $Q \in \mathbb{R}$, there is

$$\inf_{X \in \mathcal{A}(Q)} \mathcal{C}(X) = \inf_{X \in \mathcal{A}_P(Q)} \mathcal{C}(X),$$

where $\mathcal{A}_P(Q)$ is the admissible set of the pure strategy. By the hierarchy Proposition 3.3.4, we know that once the transaction-triggered price manipulation is excluded, the standard price manipulation strategy is impossible and the liquidation cost is positive. So we will first have a look at the transaction triggered price manipulation.

5.1.1 Absence of TTPM strategies on cross-impact resilience β

From the numerical example Figure 4.6 in section 4.2, we observe that if there is TTPM in zero-spread model (corresponds to the left column in the figure), TTPM is also likely to exist when $\beta < \infty$ is big but less possible to exist when β is small; when the zero-spread model excludes TTPM (corresponds to the right column), no matter what value $\beta < \infty$ is, TTPM does not present. So we conjecture that: the smaller the cross impact resilience rate β, the less profitable to apply transaction-triggered price manipulation strategies.

We start by proving that the one-side LOB model which corresponds to a $\beta = 0$ does not admit the TTPM.

Proposition 5.1.1: *Consider a cross-impact LOB model* (3.1.10) *with a constant time-varying shape function $f(t, x) = q(t)$, if the cross impact resilience rate $\beta = 0$, the model excludes transaction-triggered price manipulation.*

As the cross impact resilience rate becomes bigger, i.e. $\beta > 0$, we get the following property about TTPM between models of zero-spread (corresponds to $\beta = \infty$) and cross-impact. Note that the Proposition 5.1.2 works for any shape function $f(t, x)$ which is twice continuous differentiable, deterministic and strictly positive.

Proposition 5.1.2: *Consider a cross-impact LOB model* (3.1.10) *with a constant time-varying shape function* $f(t, x) = q(t)$*, if there is no transaction-triggered price manipulation in the LOB model with* $\beta = \infty$*, then the cross-impact LOB model with* $0 < \beta < \infty$ *does not admit transaction-triggered price manipulation.*

The economics interpretation of this cross-impact resilience rate dependent absence condition of TTPM is that: The key step for the success of the market manipulation in the real financial market is to make the market believe the misleading trading and then to create extra supply on the opposite side of the order book. If the opposite side is not responding (i.e. $\beta = 0$), it is not beneficial to trade on both sides of the order book. As the cross impact resilience rate β becomes smaller, the ability to create extra liquidity via manipulations drops.

Predoiu et al. [71] and Fruth et al. [31] are two which obtain the same results as our Proposition 5.1.1 without solving the optimal execution problem. We extend their results about the absence of TTPM. Although not explicitly stated out in their wor, Predoiu et al. [71] and Fruth et al. [31] can be regarded as our $\beta = 0$ one-side LOB model. Predoiu et al. [71] works with a time-independent shape function. Fruth et al. [31] studies a constant shape function with stochastic depth and they believe that it is the non-zero spread that precludes the TTPM. In particular, our cross impact time varying LOB model possessing a non-zero spread, admits the TTPM under some circumstances. This implies that the non-zero spread is not the reason to exclude the TTPM, but the way of modelling the two sides resilience of the order book.

5.1.2 A sufficient condition for absence of TTPM

The cost function of pure strategy under the zero-spread, cross-impact and one-side LOB model is the same by Proposition 4.1.2. Without loss of generality, given a pure buy strategy $X_P \in \mathcal{A}_P(Q)$, the cost function of pure strategy is given by

$$\mathcal{C}_P(X_P) = \int_0^T D_t^A dX_t^A + \sum_{t \le T} \frac{(\Delta X_t^A)^2}{2q(t)}. \tag{5.1.1}$$

For an arbitrary strategy $X \in \mathcal{A}(Q)$, the cross-impact cost function is

$$\begin{aligned}
\mathcal{C}^\beta(X^A, X^B) &= \int_0^T (D_t^A - L_t^A + \frac{\Delta X_t^A}{2q(t)}) dX_t^A + \int_0^T (D_t^B - L_t^B + \frac{\Delta X_t^B}{2q(t)}) dX_t^B \\
&= \underbrace{\left[\int_0^T (D_t^A - L_t^B) dX_t^A + (D_t^B - L_t^A) dX_t^B\right]}_{=E_1} \\
&\quad + \underbrace{\left[\int_0^T \frac{\Delta X_t^A}{2q(t)} dX_t^A + \int_0^T \frac{\Delta X_t^B}{2q(t)} dX_t^B\right]}_{=E_2}.
\end{aligned}$$

Since X_t^B is an increasing process in the sense that $\Delta X_t^B \ge 0$ and $dX_t^B \ge 0$ for all $t \in [0, T]$, we have

$$E_2 \ge \int_0^T \frac{\Delta X_t^A}{2q(t)} dX_t^A = \sum_{t \le T} \frac{(\Delta X_t^A)^2}{2q(t)}. \tag{5.1.2}$$

In addition, we can rewrite E_1 as

$$E_1 = \int_0^T D_t^A dX_t^A + \underbrace{\int_0^T D_t^B dX_t^B}_{=E_3} - \underbrace{\int_0^T L_t^A dX_t^A}_{=E_4} - \underbrace{\int_0^T L_t^B dX_t^B}_{=E_5}. \tag{5.1.3}$$

By definition of TTPM, if for any $X \in \mathcal{A}(Q)$ there are $X_P \in \mathcal{A}_P(Q)$ such that $\mathcal{C}^\beta(X) \ge \mathcal{C}_P(X_P)$, the cross-impact LOB model does not admit TTPM. It is sufficient to say that if $E_3 - E_4 - E_5 \ge 0$, there is no TTPM. Note however that $E_3 - E_4 - E_5 \ge 0$ is not a sharp condition. In particular, when $E4 - E5 - E6 < 0$

there might not admit TTPM if for any strategy (X^A, X^B) one has

$$\int_0^T \frac{\Delta X_t^B}{2q(t)} dX_t^B + E_3 - E_4 - E_5 > 0.$$

In the following, we illustrate an example of increasing depth function under the cross-impact LOB model and show how one could check the absence of TTPM by using the condition $E_3 - E_4 - E_5 \geq 0$.

Example 5.1.3: *Suppose that the depth function is $q(t) = e^{\alpha t}$ with $\alpha \geq 0$. Consider a strategy where shares are accumulated continuously at the positive constant rate v_1 on time interval $[0, \tau]$ and then liquidated continuously at the positive constant rate v_2 during the rest of time $(\tau, T]$ such that at the end of trading, Q shares are purchased. The time τ is one such that $v_1\tau - v_2(T - \tau) = Q$. So one gets*

$$\tau = \frac{Tv_2 + Q}{v_1 + v_2}.$$

Now we calculate E_3, E_4 and E_5 as follows:

$$\begin{aligned}
E_3 &= \int_\tau^T v_2 \int_0^t \frac{e^{-\rho(t-s)}}{e^{\alpha s}} dX_s^B dt = v_2^2 \int_\tau^T \int_\tau^t e^{-\rho t} e^{s(\rho-\alpha)} ds dt \\
&= \frac{v_2^2}{\rho - \alpha} \int_\tau^T e^{-\rho t}(e^{t(\rho-\alpha)} - e^{\tau(\rho-\alpha)}) dt \\
&= \frac{v_2^2}{\rho - \alpha} \left[\frac{e^{-\alpha\tau} - e^{-\alpha T}}{\alpha} + \frac{e^{\tau(\rho-\alpha)}}{\rho} \left(e^{-T\rho} - e^{-\tau\rho} \right) \right] \\
&= \frac{v_2^2}{\rho - \alpha} [l(\alpha) - e^{\tau(\rho-\alpha)} l(\rho)] \\
&> 0, \\
E_4 &= \int_0^\tau v_1 \int_0^t \frac{e^{-\rho(t-s)}}{e^{\alpha s}} (1 - e^{-\beta(t-s)}) dX_s^B dt \\
&= 0,
\end{aligned}$$

and

$$\begin{aligned}
E_5 &= \int_\tau^T v_2 \int_0^t e^{-\rho(t-s)-\alpha s}(1 - e^{\beta(t-s)}) dX_s^A dt \\
&= v_2 \int_\tau^T v_2 \int_0^\tau e^{-\rho(t-s)-\alpha s}(1 - e^{\beta(t-s)}) v_1 ds dt
\end{aligned}$$

$$
\begin{aligned}
&= v_1 v_2 \int_{\tau}^{T} e^{-\rho t} \int_{0}^{\tau} e^{(\rho-\alpha)s}(1 - e^{-\beta(t-s)}) ds dt \\
&= \frac{v_1 v_2}{\rho(\rho-\alpha)} \left(e^{-\tau\rho} - e^{-T\rho}\right) \left(e^{\tau(\rho-\alpha)} - 1\right) \\
&\quad - \frac{v_1 v_2}{(\rho+\beta)(\rho+\beta-\alpha)} \left(e^{-\tau(\rho+\beta)} - e^{-T(\rho+\beta)}\right) \left(e^{\tau(\rho+\beta-\alpha)} - 1\right) \\
&= v_1 v_2 l(\rho) m(\rho-\alpha) - v_1 v_2 m(\rho+\beta-\alpha) l(\rho+\beta) \\
&> 0,
\end{aligned}
$$

where functions $l(\cdot)$ and $m(\cdot)$ in the last equation are defined as

$$
l(x) := \frac{e^{-\tau x} - e^{-Tx}}{x}
$$

and

$$
m(x) := \frac{e^{\tau x} - 1}{x}.
$$

Thereafter, we obtain

$$
\begin{aligned}
E_3 - E_4 - E_5 = E_3 - E_5 \\
&= \frac{v_2}{\rho - \alpha} \left[v_2 l(\alpha) - v_2 e^{\tau(\rho-\alpha)} l(\rho) - v_1 e^{\tau(\rho-\alpha)} l(\rho) + v_1 l(\rho) \right] \\
&\quad + v_1 v_2 l(\rho+\beta) m(\rho+\beta-\alpha) := H + v_1 v_2 l(\rho+\beta) m(\rho+\beta-\alpha).
\end{aligned}
$$

Before checking the condition $E_3 - E_5 \geq 0$, we derive some properties about functions l and m. The function $m(x) > 0$ for all $\tau > 0$ and $x \in \mathbb{R}$. $l(x)$ is positive and non-increasing on $\mathbb{R}$ since its derivative is

$$
\frac{\partial l}{\partial x} = \frac{1}{x^2}[e^{-Tx}(Tx+1) - e^{-\tau x}(\tau x + 1)] = \frac{1}{x^2}[h(T) - h(\tau)],
$$

where $h(t) = e^{-tx}(tx+1)$ and $\frac{\partial h}{\partial t} = -tx^2 e^{-tx} < 0$, $\forall t > 0$.

- *Case 1: $v_1 \leq v_2$.*

Since $\frac{1}{\rho-\alpha}[l(\alpha) - e^{\tau(\rho-\alpha)}l(\rho)] > 0$, *we have*

$$\begin{aligned}
H &= \frac{v_2}{\rho-\alpha}\left[v_2 l(\alpha) - v_2 e^{\tau(\rho-\alpha)}l(\rho)\right] - \frac{v_2}{\rho-\alpha}\left[v_1 e^{\tau(\rho-\alpha)}l(\rho) - v_1 l(\rho)\right] \\
&\geq \frac{v_2 v_1}{\rho-\alpha}\left[l(\alpha) - 2e^{\tau(\rho-\alpha)}l(\rho) + l(\rho)\right] \\
&= \frac{v_2 v_1}{l(\rho)} \frac{\frac{l(\alpha)}{l(\rho)} + 1 - 2e^{\tau(\rho-\alpha)}}{\rho-\alpha} \\
&> 0,
\end{aligned}$$

if $\frac{v_2 v_1}{l(\rho)} \frac{\frac{l(\alpha)}{l(\rho)}+1-2e^{\tau(\rho-\alpha)}}{\rho-\alpha} > 0$. *This implies* $E_3 - E_5 > 0$ *since* $m(x)$ *and* $l(x)$ *are positive* $\forall \tau \in [0, T]$ *and* $x \in \mathbb{R}$.

If $H < 0$ *and* $\rho - \alpha < \rho + \beta - \alpha < 0$, *we then have* $E_3 - E_5 > 0$ *since*

$$\begin{aligned}
E_3 - E_5 &= H + v_1 v_2 l(\rho+\beta) m(\rho+\beta-\alpha) \\
&= \frac{v_2}{\rho-\alpha}\left[v_2\left(l(\alpha) - l(\rho)e^{\tau(\rho-\alpha)}\right)\right] - v_1 v_2 l(\rho) m(\rho-\alpha) \\
&\quad + v_1 v_2 l(\rho+\beta) m(\rho+\beta-\alpha) \\
&\geq \frac{v_2}{\rho-\alpha}\left[v_1\left(l(\alpha) - l(\rho)e^{\tau(\rho-\alpha)}\right)\right] - v_1 v_2 l(\rho) m(\rho-\alpha) \\
&\quad + v_1 v_2 l(\rho+\beta) m(\rho+\beta-\alpha) \\
&\geq \frac{v_1 v_2\left[l(\rho)\left(1 - e^{\tau(\rho-\alpha)}\right) + l(\rho+\beta)\left(e^{\tau(\rho+\beta-\alpha)} - 1\right)\right]}{\rho+\beta-\alpha} \\
&\quad \times \frac{v_1 v_2\left(l(\alpha) - l(\rho)e^{\tau(\rho-\alpha)}\right)}{\rho+\beta-\alpha} \\
&= \frac{v_1 v_2}{\rho+\beta-\alpha}\left[l(\alpha) + l(\rho) + l(\rho+\beta)\left(e^{\tau(\rho+\beta-\alpha)} - 1\right) - 2l(\rho)e^{\tau(\rho-\alpha)}\right] \\
&> 0.
\end{aligned}$$

- *Case 2:* $v_1 > v_2$. *There is no easier condition found to exclude the TTPM other than asking for*

$$\frac{v_2}{\rho-\alpha}\left[v_2 l(\alpha) - v_2 e^{\tau(\rho-\alpha)}l(\rho)\right] - \frac{v_2}{\rho-\alpha}\left[v_1 e^{\tau(\rho-\alpha)}l(\rho) - v_1 l(\rho)\right] > 0$$

or

$$\frac{v_2}{\rho-\alpha}\left[v_2 l(\alpha) - v_2 e^{\tau(\rho-\alpha)} l(\rho) - v_1 e^{\tau(\rho-\alpha)} l(\rho) + v_1 l(\rho)\right] + l(\rho+\beta) m(\rho+\beta-\alpha) \geq 0.$$

5.1.3 Absence of TTPM strategies with time-independent shape function

In this section only, we do not assume a constant shape function. A time-independent shape function takes the form of $f(t,x) = f(x)$.

Some zero-spread LOB models (i.e. $\beta = \infty$), such as Alfonsi et al. [6], Alfonsi et al. [5] and Alfonsi et al. [4], work under assumption of time-independent shape function. Their results are that the time independent shape limit order book model with zero spread precludes the TTPM.

Here we extend their results to the cross impact modelling where the opposite side resilience rate β is finite. We then claim that the TTPM is impossible under the cross-impact LOB model in case of time-independent shape function. Suppose there is a minimiser strategy $X^{*,\infty}$ for zero-spread LOB cost function and a minimiser strategy $X^{*,\beta}$ for cross-impact LOB cost function, Proposition 5.1.2 implies that if $X^{*,\infty}$ is a pure strategy, then the optimal strategy under cross impact model $X^{*,\beta}$ is also a pure strategy. That is to say, the TTPM is absent under all three LOB modes when the shape function is time-independent.

5.1.4 Absence of TTPM strategies on trading frequency

An interesting property of the optimal strategy under the assumption of time independent shape function is that the optimal strategy consists in a sequence of market orders that consume exactly that amount of shares by which the LOB has recovered since the proceeding market order due to the resilience effect. If this is true for time-varying shape function as well, the TTPM can be excluded by checking the resilience of the order book. We then find this relationship does not hold if the shape function is time-dependent. We will show one example that $E_{t_i+} - E_{t_{i+1}} \neq \xi_{i+1}$ under the zero-spread LOB model.

Recall that the shape function is given by $f(t,x) = q(t)$. Given the trading interval $[0,T]$, the discrete time trading times are $t_i := t_0 + i\tau$ for $i = 0, ..., N$ where $\tau = \frac{T}{N}$. Let us denote by $a_i = e^{-\rho(t_i - t_{i-1})}$ and $\hat{a}_i = a_i \frac{q(t_i)}{q(t_{i-1})}$ for $1 \leq i \leq N$. Note that t_i and $\hat{a}_i$ are functions of N.

In zero-spread LOB model framework, the optimal execution problem admits a unique optimal strategy $\xi^* \in \mathcal{A}(Q)$ which is explicitly given by

$$\begin{cases} \xi_0^* = \frac{Q}{K} q(t_0) \frac{1-\hat{a}_1}{1-a_1\hat{a}_1}, \\ \xi_i^* = \frac{Q}{K} q(t_i) \left[\frac{a_i}{1-a_i\hat{a}_i}(\hat{a}_i - 1) + \frac{1-\hat{a}_{i+1}}{1-a_{i+1}\hat{a}_{i+1}} \right], \quad 1 \leq i \leq N-1 \\ \xi_N^* = \frac{Q}{K} q(t_N) \frac{1-a_N}{1-a_N\hat{a}_N}, \end{cases} \tag{5.1.4}$$

where

$$K = \frac{q(t_N)(1-2a_N) + q(t_{N-1})}{\alpha_N(1-a_N\hat{a}_N)} + \sum_{i=0}^{N-2} q(t_i) \frac{(1-\hat{a}_{i+1})^2}{(1-a_{i+1}\hat{a}_{i+1})}.$$

We find that the optimal strategy is not perfectly using the resilient amount of the order book during each small trading interval. The resilience amount over time interval (t_i, t_{i+1}) is

$$E_{t_i+} - E_{t_{i+1}} = \frac{Q}{K} \frac{1-\hat{a}_{i+1}}{1-a_{i+1}\hat{a}_{i+1}} (q(t_i) - a_{i+1} q(t_{i+1})).$$

The optimal trading size during this same interval is

$$\xi_{i+1}^* = q(t_{i+1}) \frac{1-\hat{a}_{i+2}}{1-a_{i+2}\hat{a}_{i+2}} - q(t_{i+1}) a_{i+1} \frac{1-\hat{a}_{i+1}}{1-a_{i+1}\hat{a}_{i+1}} \neq E_{t_i+} - E_{t_{i+1}}.$$

Furthermore, one need to note that the length of each small trading interval still has an affect on the presence of transaction-triggered price manipulation. We will have a look at two examples under the framework of zero-spread LOB model

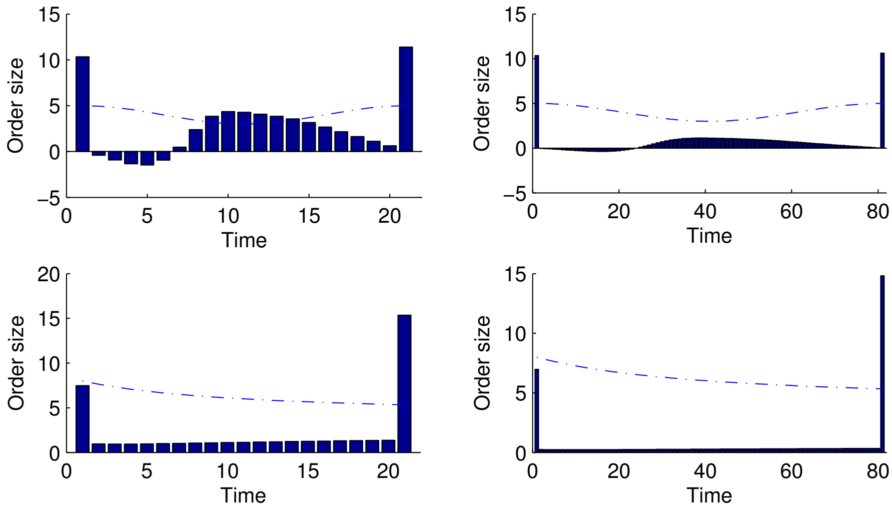


Figure 5.1: The optimal execution strategies for buying 50 shares on a regular time grid with $N = 20$ (plots in left column) and $N = 80$ (plots in right column). For each sub-figure, fix $T = 1$, $\rho = 2$ and $\beta \to \infty$. Both plots in the first row illustrate the strategies with $q(t) = 4 + \cos(2\pi t)$ (plotted in dashed lines) and both plots in the second row illustrate the strategies with $q(t) = 4 + \frac{2}{0.5+t}$ (plotted in dashed lines).

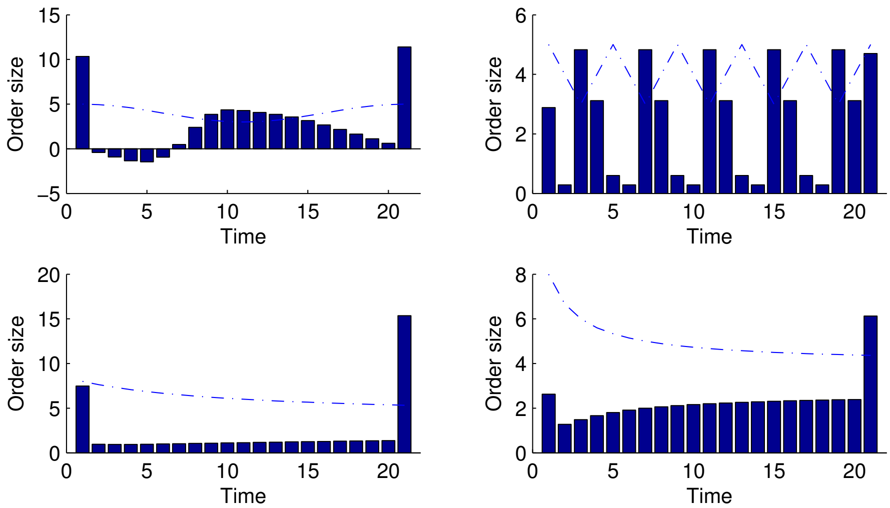


Figure 5.2: The optimal execution strategies for buying 50 shares on a regular time grid with $T = 1$ (plots in left column) and $T = 5$ (plots in right column). For each sub-figure, fix $N = 20$, $\rho = 2$ and $\beta \to \infty$. Both plots in the first row illustrate the strategies with $q(t) = 4 + \cos(2\pi t)$ (plotted in dashed lines) and both plots in the second row illustrate the strategies with $q(t) = 4 + \frac{2}{0.5+t}$ (plotted in dashed lines).

with time-varying, constant shape function. In Figure 5.1 we change each trading times by fixing the trading time interval T and changing the trading frequency from 20 to 80. The lump orders at the beginning and the end of T have no change, but the intermediate orders becomes smaller as N increases. In Figure 5.2 we fix the trading frequency $N = 20$ and change the trading time interval T from 1 to 5. Focusing on the upper row plots, when T is increasing, the optimal strategy becomes a pure strategy. For the models excluding the TTPM (lower row plots), the sizes of the first and last order become smaller and the child orders are more evenly distributed.

5.2 Absence and presence of price manipulation strategies (PMS)

Recall that a market impact model dose not admit price manipulation strategy (PMS) if $\inf_{X\in\mathcal{A}(0)} \mathcal{C}(X) \geq 0$ where $\mathcal{A}(0)$ is the set of the round-trip strategies which are to buy the total amount of zero shares. From the analysis of TTPM absence section 5.1 and irregularity hierarchy relationship Proposition 3.3.4, all three market irregularities can be excluded under the one-side LOB model. Thus we only study the cross-impact and zero-spread models in this section for price manipulation strategy and the next section for positive liquidation cost.

5.2.1 Absence of PM strategy on β

Proposition 4.1.2 shows that given a trading strategy, the cost under the zero-spread model is a lower bound of the cost under cross-impact model, and the cost under the one-side model is an upper bound of the cross-impact cost. In terms of the price manipulation, this proposition implies that the profitable round trips under the cross-impact model also make non-negative costs under the zero-spread model. Note that this PMS property holds even for a general shape function.

Proposition 5.2.1: *Consider $\mathcal{C}^\beta$ the cross-impact cost function (3.1.10) and $\mathcal{C}^\infty$*

the zero-spread cost function (3.2.2), if there is round trip strategy $X \in \mathcal{A}(0)$ *such that* $\mathcal{C}^{\beta}(X) < 0$*, we have* $\mathcal{C}^{\infty}(X) < 0$.

In other words, if PMS is impossible in zero-spread LOB model, then there is no PMS in cross-impact LOB model. However, in general, the converse does not hold. We will present in Example 5.2.2 the possibility of $\mathcal{C}^{\infty} < 0 \leq \mathcal{C}^{\beta}$.

Example 5.2.2: *Consider a round trip* $X = (\Delta X^A_{t_i}, \Delta X^B_{t_j}) = (m, m)$ *with* $t_i < t_j < T$*. The cross impact cost function of this strategy is given by*

$$\begin{aligned}
\mathcal{C}^{\beta}(X) &= \left(\frac{m^2}{2q(t_i)} + \frac{m^2}{2q(t_j)}\right) - \frac{m^2 e^{-\rho(t_j - t_i)}}{q(t_i)}\left(1 - e^{-\beta(t_j - t_i)}\right) \\
&= m^2\left[\left(\frac{1}{2q(t_i)} + \frac{1}{2q(t_j)}\right) - \frac{a}{q(t_i)}(1-c)\right] \\
&= \frac{m^2}{2q(t_i)}(b + 1 - 2a + 2ac),
\end{aligned}$$

where $a = e^{-\rho(t_j - t_i)}$, $b = \frac{q(t_i)}{q(t_j)}$ *and* $c = e^{-\beta(t_j - t_i)}$.

Applying the same strategy $X = (m, m)$ *to the zero-spread cost function (5.0.3), one obtains the zero-spread cost function*

$$\begin{aligned}
\mathcal{C}^{\infty}(X) &= m^2\left[\left(\frac{1}{2q(t_i)} + \frac{1}{2q(t_j)}\right) - \frac{a}{q(t_i)}\right] \\
&= \frac{m^2}{2q(t_i)}(b + 1 - 2a).
\end{aligned}$$

If ρ*,* β *and* α *are such that* $\beta \leq \frac{1}{t_j - t_i}\ln\frac{2a}{-(b+1-2a)}$ *and* $b + 1 - 2a < 0$*, one obtains* $\frac{m^2}{2q(t_i)}(b+1-2a) < 0 \leq \frac{m^2}{2q(t_i)}(b+1-2a+2ac)$*. An economics interpretation of the condition* $\beta \leq \frac{1}{t_j - t_i}\ln\frac{2a}{-(b+1-2a)}$ *is that: As the cross-impact factor* β *becomes smaller, from* ∞ *to less than* $\frac{1}{t_j - t_i}\ln\frac{2a}{-(b+1-2a)}$*, the opposite side resilience becomes slower. This results in smaller opposite side price recovery, and then the profits of the round trip is smaller, e.g. after a buy order, one might sell at a lower price if the opposite side resilience is smaller.*

5.2.2 Conditions for absence of PM strategies

In this section, we will first present two necessary conditions for absence of PMS under the cross-impact model.

Proposition 5.2.3: *Consider a cross-impact LOB model* (3.1.10) *with a constant time-varying shape function* $f(t, x) = q(t)$*, if the cross-impact LOB model does not admit the price manipulation strategy, then for any* $s < t$ *the depth function satisfies*

$$\frac{q(s)}{q(t)} \geq 2e^{-\rho(t-s)} - 1 - 2e^{-(\rho+\beta)(t-s)}. \tag{5.2.1}$$

Corollary 5.2.4: *Consider a cross-impact LOB model* (3.1.10) *with a constant time-varying shape function* $f(t, x) = q(t)$*, if the price manipulation strategy is excluded in the cross-impact LOB model, then the cross-impact resilience rate satisfies* $\beta(t - s) \leq \ln 2$ *for any* $s < t$*.*

An extreme example of this Corollary 5.2.4 is that if $\beta = 0$, we are then in the one-side LOB world, and we know PMS is not possible.

Next, let us consider the corresponding results for zero-spread model. By sending the cross impact resilience rate β to ∞, one obtains the necessary condition for LOB model with zero spread, which is given by

$$\frac{q(s)}{q(t)} \geq 2e^{-\rho(t-s)} - 1. \tag{5.2.2}$$

In general, condition (5.2.1) and condition (5.2.2) are not sufficient. We will present in the following example that for a depth function $q(t)$ satisfying condition (5.2.2), the zero-spread LOB model still admits price manipulation strategy.

Example 5.2.5: *Let us consider a three-trade round trip strategy under zero spread LOB model (i.e. the cross-impact resilience rate* $\beta = \infty$*). Trading times are* $\mathbb{T} = \{t_0, t_1, t_2\}$ *and total trading size is* $Q = 0$*.*

Given a depth function $q(t) = (1+t)^2$ *and a same side resilience rate* $\rho = 0.4$*,* *one has*

$$\frac{q(t_i)}{q(t_j)} = \left(\frac{1+t_i}{1+t_j}\right)^2 = 1 - \frac{2(t_j - t_i)}{1+t_j} + \left(\frac{t_j - t_i}{1+t_j}\right)^2 > 1 - \frac{2(t_j - t_i)}{1+t_j}$$

and

$$1 - \frac{(t_j - t_i)}{1+t_j} \geq 1 + \frac{t_j}{1+t_j} > 1 \geq e^{-\rho(t_j - t_i)}.$$

Thus, given $1 - \frac{(t_j - t_i)}{1+t_j} > e^{-\rho(t_j - t_i)}$, *we obtain that*

$$1 - \frac{2(t_j - t_i)}{1+t_j} \geq 2e^{-\rho(t_j - t_i)} - 1,$$

which implies that condition (5.2.1) is satisfied. Now, let us look at the cost function of round trip $(x_0, x_1, -x_0 - x_1)$, *which is in the form of*

$$\begin{aligned}
\mathcal{C}^\infty(x_0, x_1, -x_0 - x_1) = -(x_0 + x_1)&\left(\frac{1}{9}(-x_0 - x_1) + 0.4493x_0 + 0.1676x_1\right) \\
&+ x_0(0.4493(-x_0 - x_1) + x_0 + 0.6703x_1) \\
&+ x_1(0.1676(-x_0 - x_1) + 0.6703x_0 + x_1/4).
\end{aligned}$$

Then we solve the inequality $\mathcal{C}^\infty(x_0, x_1, -x_0 - x_1) < 0$ *and the solutions are*

$$x_0 < 0 \qquad \text{and} \qquad -6.34906x_0 - 5.66648|x_0| < x_1 < -6.34906x_0 + 5.66648|x_0|$$

or

$$x_0 > 0 \qquad \text{and} \qquad -6.34906x_0 - 5.66648|x_0| < x_1 < -6.34906x_0 + 5.66648|x_0|.$$

5.3 Conditions for the positive liquidation cost (PLC)

Recall first that a market impact model has positive liquidation costs (PLC) if for $\forall Q \in \mathbb{R}$, there is

$$\inf_{X \in \mathcal{A}(Q)} \mathcal{C}(X) \geq 0.$$

First, we try to derive a discrete-time necessary and sufficient condition for PLC. Recall the cross-impact cost function (4.2.2) with constant shape function

$$C^\beta(\Theta^A, \Theta^B) = \frac{1}{2}\langle z, Hz\rangle, \tag{5.3.1}$$

where

$$\begin{aligned}
&\Theta^A \in \mathcal{A}_N(Q), && \Theta^{\in}\mathcal{A}_N(Q), && \\
&a_{i,j} := \frac{e^{-\rho(t_j-t_i)}}{q(t_i)}, && \tilde{a}_{i,j} := \frac{e^{-(\rho+\beta)(t_j-t_i)}}{q(t_i)}, && \\
&A := a_{i,j}\mathbb{1}_{\{i<j\}}, && \tilde{A} := \tilde{a}_{i,j}\mathbb{1}_{\{i<j\}}, && \bar{A} := a_{i,j}\mathbb{1}_{\{i<j\}} + \frac{a_{i,j}}{2}\mathbb{1}_{\{i=j\}}, \\
&B = \frac{1}{2}(\bar{A}^T + \bar{A}), && D = A - \tilde{A}, &&
\end{aligned}$$

and

$$z := (\Theta^A, \Theta^B), \qquad M := \begin{pmatrix} B & -D \\ -D & B \end{pmatrix}, \quad H = M + M^T.$$

In this case, we have positive liquidation cost if and only if the depth function $q(t)$ is such that the matrix H is co-positive. A discussion of criteria for co-positive matrices can be found in Väliaho [84] and Hiriart-Urruty and Seeger [46] and the references therein.

In general, this co-positivity matrix condition for PLC is not very practical since it is not easy to check the co-positivity. We then derive a necessary condition for PLC under cross-impact LOB model.

Proposition 5.3.1: *Consider a cross-impact LOB model* (3.1.10) *with a constant time-varying shape function $f(t,x) = q(t)$, if the cross impact LOB model has positive expected liquidation cost, then for any $s < t$ the depth function satisfies*

$$\frac{q(s)}{q(t)} \geq e^{-2\rho(t-s)}(1 - e^{-\beta(t-s)})^2. \tag{5.3.2}$$

Next, we will derive a necessary and sufficient condition for PLC under zero-spread LOB model. Note that this condition (5.3.3) is a sufficient condition for PLC under cross-impact model by Proposition 4.1.2. When the cross-impact resilience rate $\beta \to \infty$, one obtains the cost function of the zero-spread LOB model

$$\mathcal{C}^{\infty}(\xi) = \frac{1}{2}\langle \xi, \Lambda \xi \rangle,$$

where $\xi \in \mathcal{A}_N^{\infty}(Q)$ and $\Lambda_{i,j} = \frac{e^{-\rho|t_j - t_i|}}{q(t_i \wedge t_j)}$ for $0 \leq i, j \leq N$ with $(m \wedge n)$ taking minimum between m and n.

Proposition 5.3.2: *Consider a zero-spread LOB model* (3.2.2) *with a constant time-varying shape function $f(t,x) = q(t)$, the zero spread LOB model has positive liquidation costs if and only if $q(t)$ satisfies*

$$q(s) \geq q(t)e^{-2\rho(t-s)}, \tag{5.3.3}$$

for any $s < t$.

5.4 Examples of models admitting PM strategies

In this section, we will show examples of price impact dynamics $A - A^0$ and $B - B^0$ generated by three LOB models which admit PMS. The three LOB models are: the zero-spread increasing depth LOB model, the zero-spread reverting depth LOB model and the cross-impact increasing LOB model. The objective is to understand that how a price manipulation strategy works under different trading times, depth function $q(t)$ and cross resilience rate β.

The base assumptions for all three examples are: a round trip strategy with two discrete trades is considered, i.e. the buy strategy is $X^A = (0, ..., 0, \Delta X^A_{t_1}, 0, ..., 0)$ and the sell strategy is $X^B = (0, ..., 0, \Delta X^B_{t_2}, 0, ..., 0)$ with $t_1 < t_2$ and $\Delta X^A_{t_1} = \Delta X^B_{t_2} := m$. Take trading time interval $T = 1$. Thereafter, the cost function of such a round trip can be expressed as a function of trading time t_1, t_2, the depth function

$q(t)$ and the cross-impact resilience rate β

$$\mathcal{C}(t_1, t_2, \beta, q(t)) = \frac{m^2 \left(\frac{q(t_1)}{q(t_2)} + 1 + 2e^{-(\rho+\beta)(t_2-t_1)} - 2e^{\rho(t_2-t_1)} \right)}{2q(t_1)}. \tag{5.4.1}$$

Example 5.4.1 (Zero-spread and increasing depth LOB model)**:** *Consider an exponential depth function $q(t) = q(0)e^{\alpha t}$. First, we investigate that how the costs change against trading time t_1 and t_2. When t_1 is fixed, like in Figure 5.3 where we set $t_1 = 0.2$, we see that placing the sell order at different t_2 might lead to negative costs, which means this particular LOB model admits PMS. In other words, in order to create negative costs the trader should submit the sell order between time $(0.2, 0.7)$.*

Next, we look at how the ask side and bid side price impact dynamics behave. Figure 5.4 and Figure 5.5 illustrate two price impact dynamic examples when trading times are $t_1 = 0.2$, $t_2 = 0.4$ and $t_1 = 0.2$, $t_2 = 0.8$ respectively.

The price impact dynamics can be described as: before first buy order at t_1, there is no price impact incurred on both sides. At time t_1 a buy order of size m pushes the best ask price up and makes a hole in the order book. This is reflected by the vertical line at time t_1 in plots of $A - A^0$. The bid side refills the hole at an infinite speed (i.e. $\beta \to \infty$), which is illustrated by the vertical line at t_1 in plots of $B - B^0$. During time (t_1, t_2), price impact decays on both sides. At time t_2, another sell order further drags back the best bid price and makes a hole in the order book. This is reflected by the vertical line at t_2 in plots of $B - B^0$. The infinite resilience of the ask side refills the hole at once, which is illustrated by the vertical line at t_2 in plots of $A - A^0$.

Because of the infinite cross impact resilience, the ask side dynamic and the bid side dynamic are the same. The position of two midpoints on the two vertical lines in each plot shows whether the cost of this round trip with trading time t_1, t_2 is negative or non-negative. In Figure 5.4, midpoint $+$ is lower than the midpoint $\triangle$ implies the average buy price is lower than the average sell price. Thus, the total

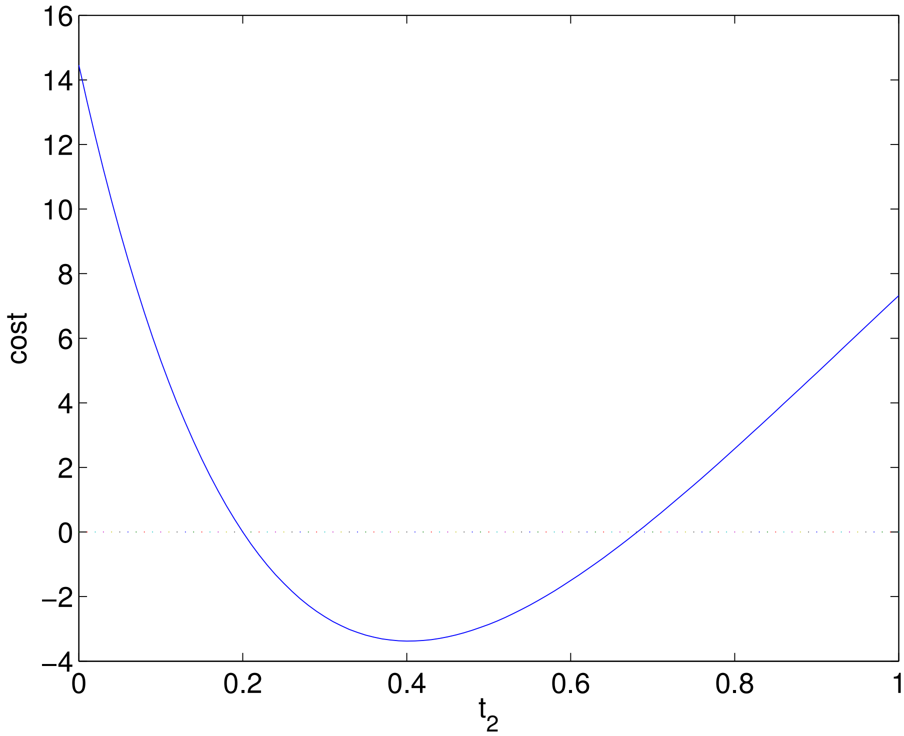


Figure 5.3: The change of the cost function against t_2 when $t_1 = 0.2$, $q(t) = q(0)e^{\alpha t}$ with $q(0) = 2$, $\alpha = 3$, $\rho = 1$, $\beta \to \infty$ and $m = 50/3$.

cost of $t_1 = 0.2$*,* $t_2 = 0.4$ *strategy is negative. While in Figure 5.5, midpoint + is higher than the midpoint* $\triangle$ *implies the total cost of* $t_1 = 0.2$*,* $t_2 = 0.8$ *strategy is positive.*

Example 5.4.2 (Zero-spread and reverting depth LOB model)**:** *Consider a reverting depth function* $q(t) = 2 - \cos(2\pi t)$*. Similarly, Figure 5.12 shows how the cost (5.4.1) changes with respect to* t_2 *for fixed time* $t_1 = 0.2$*. We also observe the price manipulation opportunities where the cost function is below zero for* $t_2 \in (0.2, 0.46)$.

The interpretation of price impact dynamics in Figure 5.6 and Figure 5.7 is the same as in Example 5.4.1. The way to detect PMS is the same as in Example 5.4.1 too. The difference between this LOB model and the one in Example 5.4.1 is the depth function. At the same time $t_2 = 0.8$*, the different depths* $q(t_2)$ *generate two different price impact dynamics as shown in Figure 5.5 and Figure 5.7.*

Example 5.4.3 (Cross-impact increasing depth LOB model)**:** *In this example, it is assumed that the depth function takes the form* $q(t) = q(0)e^{\alpha t}$*. Figure 5.8 illustrates how the cross-impact cost are affected by the trading time* t_2 *when* t_1 *is fixed and by the cross-impact resilience rate* β*. As the cross impact resilience rate* β *becomes smaller, the negative cost time interval becomes smaller and the profits* $-\mathcal{C}^{\beta}$ *is smaller. This observation coincides with Corollary 5.2.4 and Proposition 5.2.1 about the relationship between irregularity and cross-impact resilience rate.*

Three cases of price impact dynamics are presented in Figure 5.9, Figure 5.10 and Figure 5.11 respectively. The two trading times are fixed to be $t_1 = 0.2$ *and* $t_2 = 0.5$*. Compared with the price impact dynamics in Example 5.4.1, for the bid side price impact, after the first trade at time* t_1 *there is no instantaneously refill in the sense that it is not a vertical line at* t_1*. The transient opposite resilience is reflected in the ask side price dynamics at time* t_2 *as well.*

Compare the price impact dynamics in Figure 5.9 and Figure 5.10. They are distinguished by different cross impact resilience rate β*. Without loss of generality,*

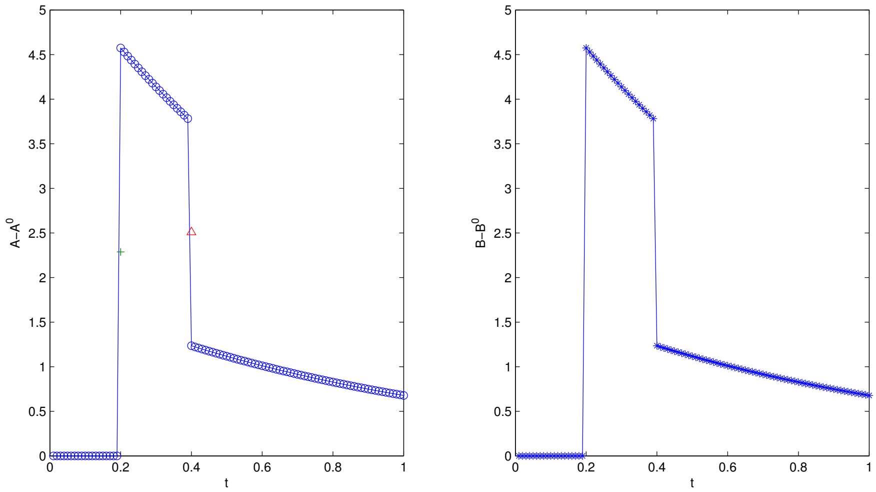


Figure 5.4: The ask side price impact (left) and the bid side price impact (right) dynamics on time interval $[0, 1]$ with $t_1 = 0.2$ and $t_2 = 0.4$.
For each plot, $\rho = 1$, $\beta \to \infty$, $q(t) = q(0)e^{\alpha t}$ with $q(0) = 2$ and $\alpha = 3$, $m = 50/3$. The point $+$ is the midpoint of the vertical line at time t_1. The point $\triangle$ is the midpoint of the vertical line at time t_2.

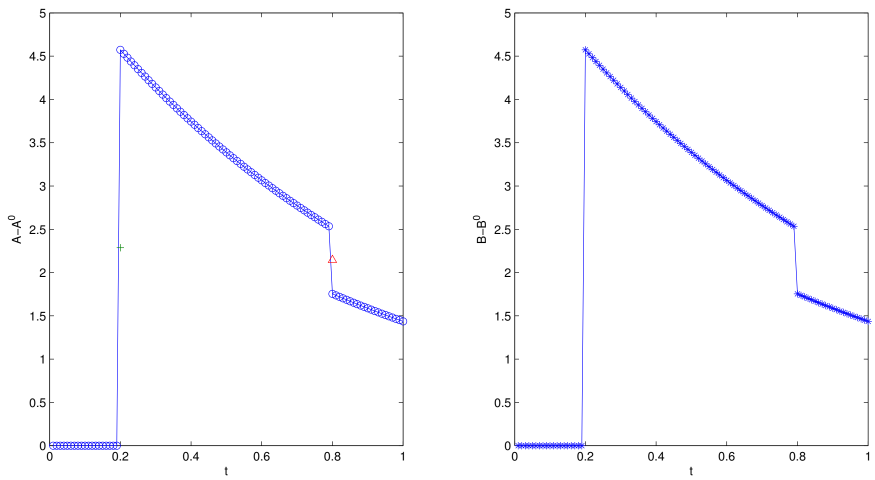


Figure 5.5: The ask side price impact (left) and the bid side price impact (right) dynamics on time interval $[0, 1]$ with $t_1 = 0.2$ and $t_2 = 0.8$.
For each plot, $\rho = 1$, $\beta \to \infty$, $q(t) = q(0)e^{\alpha t}$ with $q(0) = 2$ and $\alpha = 3$, $m = 50/3$. The point $+$ is the midpoint of the vertical line at time t_1. The point $\triangle$ is the midpoint of the vertical line at time t_2.

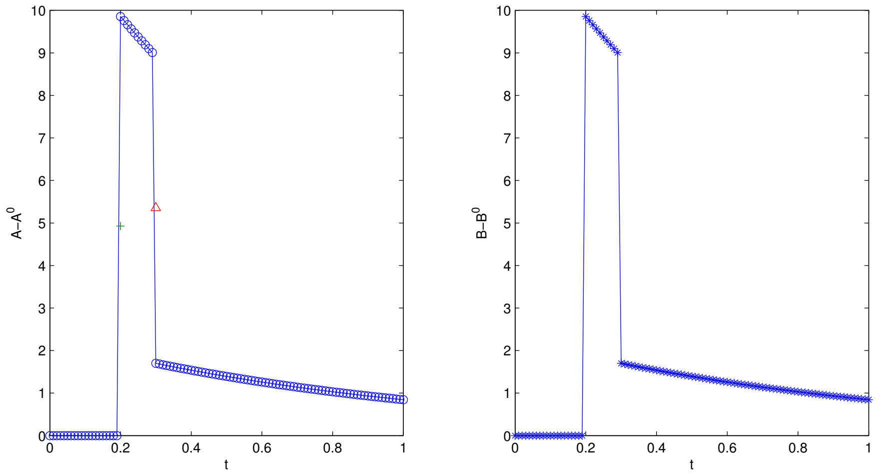


Figure 5.6: The ask side price impact (left) and the bid side price impact (right) dynamics on time interval $[0, 1]$ with $t_1 = 0.2$ and $t_2 = 0.3$.
For each plot, $\rho = 1$, $\beta \to \infty$, $q(t) = 2 - \cos(2\pi t)$, $m = 50/3$. The point $+$ is the midpoint of the vertical line at time t_1. The point $\triangle$ is the midpoint of the vertical line at time t_2.

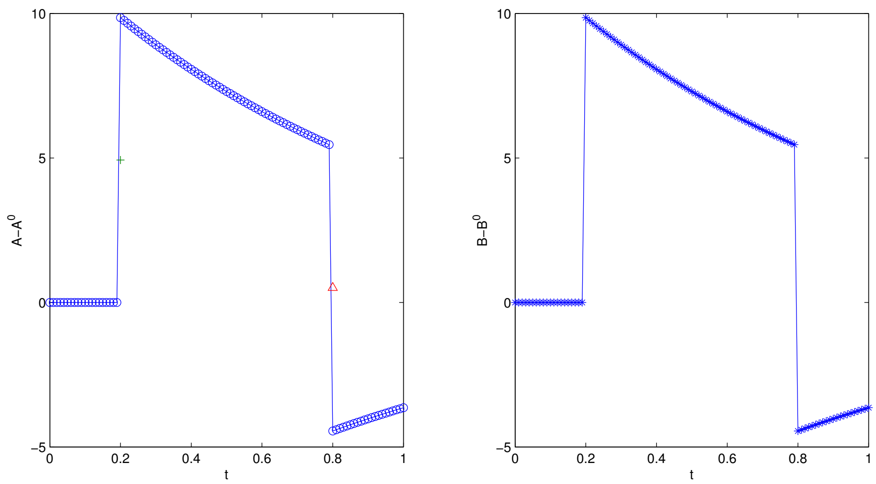


Figure 5.7: The ask side price impact (left) and the bid side price impact (right) dynamics on time interval $[0, 1]$ with $t_1 = 0.2$ and $t_2 = 0.8$.
For each plot, $\rho = 1$, $\beta \to \infty$, $q(t) = 2 - \cos(2\pi t)$, $m = 50/3$. The point $+$ is the midpoint of the vertical line at time t_1. The point $\triangle$ is the midpoint of the vertical line at time t_2.

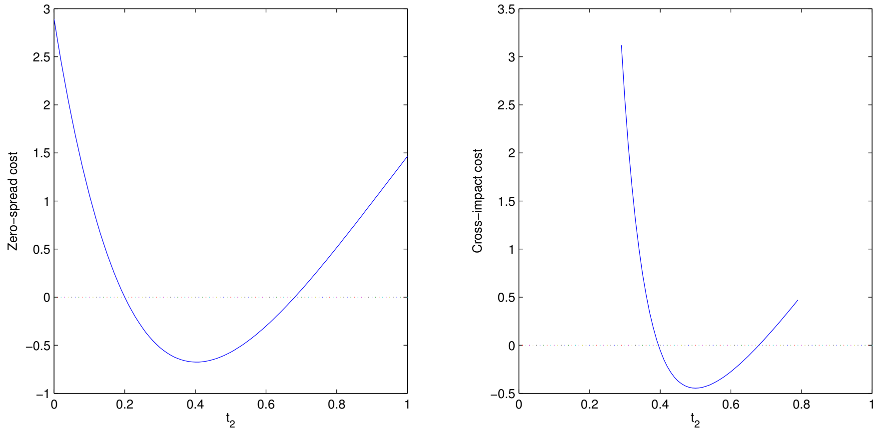


Figure 5.8: The change of the cost function against t_2 with $\beta \to \infty$ (left) and $\beta = 15$ (right) when $t_1 = 0.2$. For each plots, $\rho = 1$, $m = 50/3$, $q(t) = q(0)e^{\alpha t}$ with $q(0) = 10$ and $\alpha = 3$.

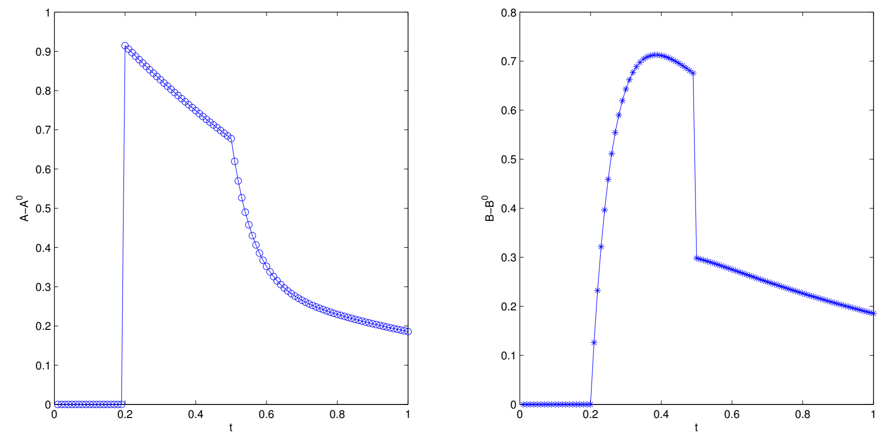


Figure 5.9: The ask side price impact (left) and the bid side price impact (right) dynamics on time interval $[0, 1]$ with $\rho = 1$ and $\beta = 15$.
For each plot, $q(t) = q(0)e^{\alpha t}$ with $q(0) = 10$ and $\alpha = 3$, $m = 50/3$, $t_1 = 0.2$ and $t_2 = 0.5$.

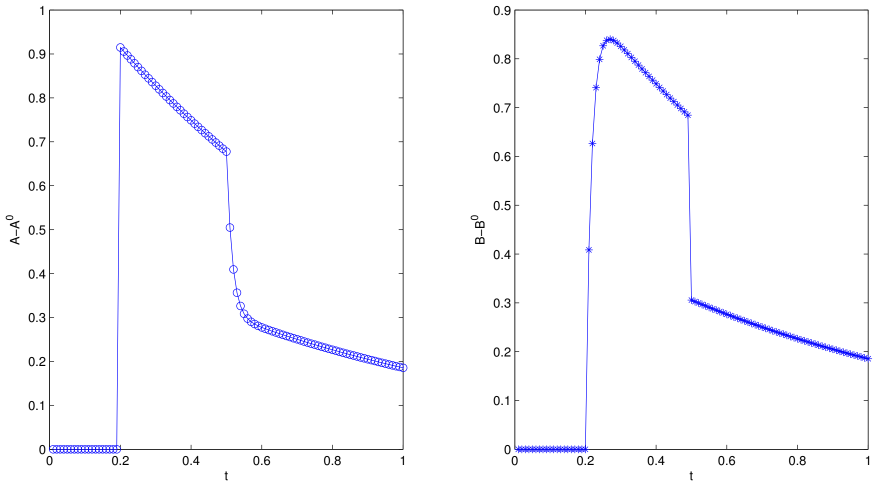


Figure 5.10: The ask side price impact (left) and the bid side price impact (right) dynamics on time interval $[0, 1]$ with $\rho = 1$ and $\beta = 60$.
For each plot, $q(t) = q(0)e^{\alpha t}$ with $q(0) = 10$ and $\alpha = 3$, $m = 50/3$, $t_1 = 0.2$ and $t_2 = 0.5$.

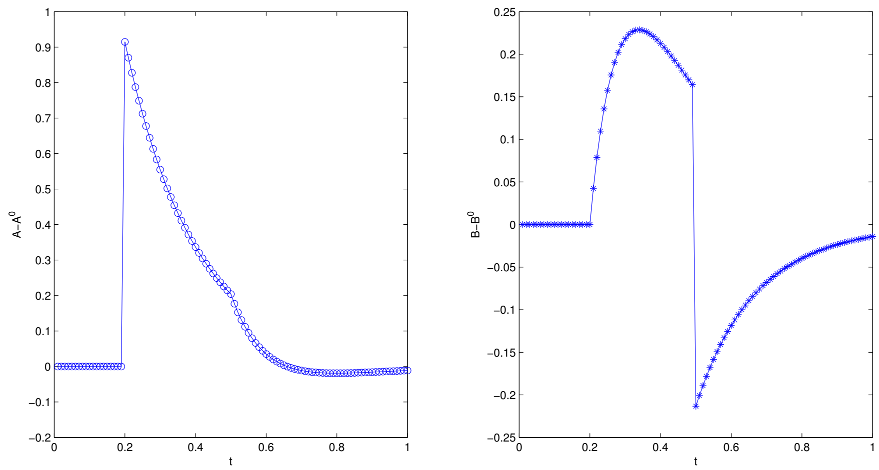


Figure 5.11: The ask side price impact (left) and the bid side price impact (right) dynamics on time interval $[0, 1]$ with $\rho = 5$ and $\beta = 15$.
For each plot, $q(t) = q(0)e^{\alpha t}$ with $q(0) = 10$ and $\alpha = 3$, $m = 50/3$, $t_1 = 0.2$ and $t_2 = 0.5$.

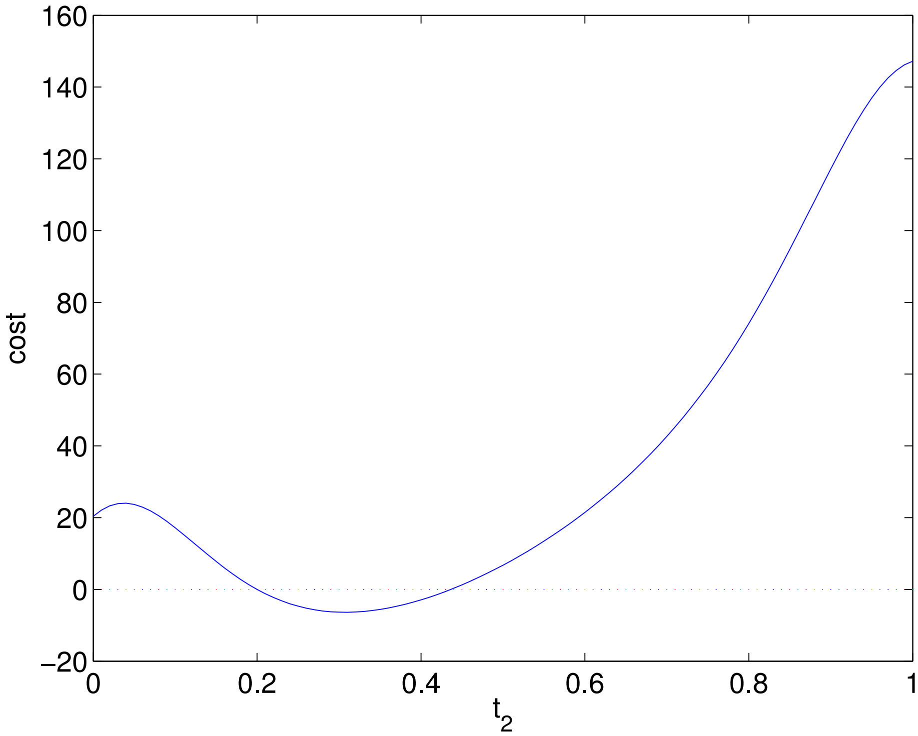


Figure 5.12: The change of the cost function against t_2 when $t_1 = 0.2$, $\rho = 1$, $\beta \to \infty$, $q(t) = 2 - \cos(2\pi t)$ and $m = 50/3$.

we look at the bid side price impact plot. As β become larger, the opposite side refills the hole faster. This is reflected in the plot by the shorter time that it takes to reach the maximum bid side price impact, whose economics interpretation is the minimum bid-ask spread after the trade at t_1. The same analysis applies to ask side price impact during time $(t_2, 1]$.

Compare the price impact dynamics in Figure 5.9 and Figure 5.11. There are different same side resilience rate ρ in these two figures. Without loss of generality, let us focus on the price impact dynamics on time interval $(t_1,\ t_2)$. After a buy market order at t_1, the ask side (i.e. same side of buy market order) decays faster if the ρ is bigger. Correspondingly, on the same time interval (t_1, t_2) at the bid side, the cross price impact incurred by the buy order at t_1 is smaller than that in Figure 5.9.

Recall that one can see the profitability of round trip strategies by investigating the levels of average buy and sell prices. In order to compare the average buy and sell prices, we do the phase portrait of the coupled system of best ask and best bid in Figure 5.13. The trade at time t_1 corresponds to the jump from 0 to 0.92 along the $A - A^0$ axis. The sell order at time t_2 is the $B - B^0$ jump from 0.69 to 0.3 in the opposite direction of $B - B^0$ axis. The point $+$ is the projection of the midpoint of $A - A^0$ jump onto the line $x = y$. The point $\triangle$ is the projection of the midpoint of $B - B^0$ jump onto the line $x = y$. Similar to the analysis in Example 5.4.1, if the point $\triangle$ is higher than the point $+$, it implies the strategy is a PMS.

5.5 Proofs

Proof of Proposition 5.1.1. Without loss of generality, we consider a purchase program. Given $\beta = 0$ in the sense that the one-side LOB model is considered, the cost function (5.0.2) can be rewritten as

$$\mathcal{C}^0(X^A, X^B) = \int_0^T \left(D_t^A + \frac{\Delta X_t^A}{2q(t)} \right) dX_t^A + \int_0^T \left(D_t^B + \frac{\Delta X_t^B}{2q(t)} \right) dX_t^B.$$

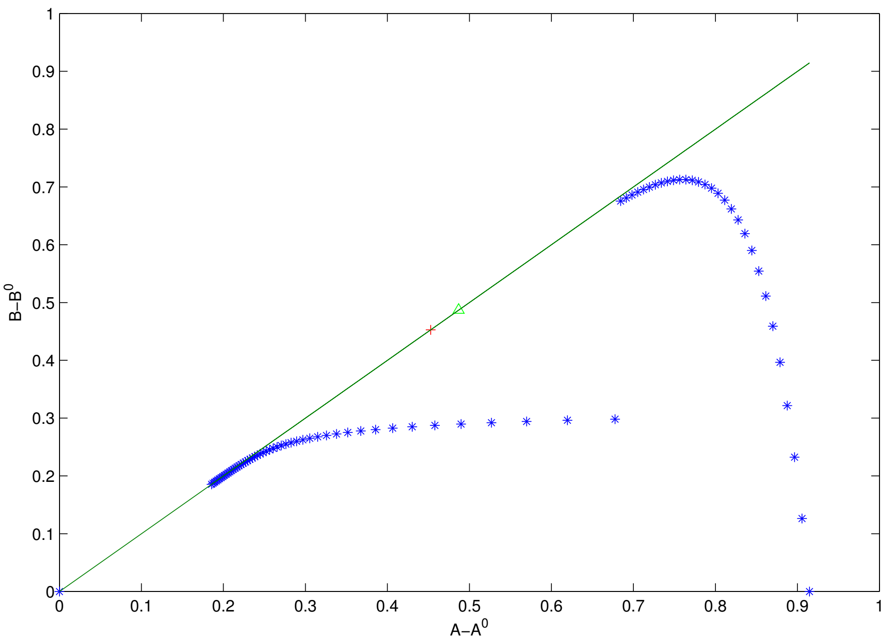


Figure 5.13: The phase portrait of the coupled system of ask side price impact against bid side price impact. $\rho = 1$, $\beta = 15$, $q(t) = q(0)e^{\alpha t}$ with $q(0) = 10$ and $\alpha = 3$, $m = 50/3$, $t_1 = 0.2$ and $t_2 = 0.5$.

Since $D_t^B \geq 0$ and X_t^B is an increasing process along time, one obtains

$$\mathcal{C}^0(X^A, X^B) \geq \int_0^T \left(D_t^A + \frac{\Delta X_t^A}{2q(t)} \right) dX_t^A. \tag{5.5.1}$$

The right hand side of inequality (5.5.1) is the cost function of a pure buy strategy $(X_t^A, 0)$ to buy a total amount of $X_{T+}^A > Q$ since $Q = X_{T+}^A - X_{T+}^B$. If we replace the pure buy strategy X_t^A by $\min\{X_t^A, Q\}$ for all $t \in [0, T]$, we obtain a feasible pure buy strategy in $\mathcal{A}_P(Q)$ whose total cost is less than or equal to $\int_0^T \left(D_t^A + \frac{\Delta X_t^A}{2q(t)} \right) dX_t^A$. Therefore, the transaction-triggered price manipulation is excluded. □

Proof of Proposition 5.1.2. Given that there is no TTPM under the zero-spread LOB model, by the definition of transaction-triggered price manipulation, one has the following relationship

$$\inf_{X \in \mathcal{A}(Q)} \mathcal{C}^\infty(X) = \inf_{X \in \mathcal{A}_P(Q)} \mathcal{C}^\infty(X).$$

Next, let us consider the minimum costs $\mathcal{C}^{*,\infty} := \inf_{X \in \mathcal{A}(Q)} \mathcal{C}^\infty(X)$ and $\mathcal{C}^{*,\beta} := \inf_{X \in \mathcal{A}(Q)} \mathcal{C}^\beta(X)$. By Corollary 4.1.3, we have

$$\mathcal{C}^{*,\infty} \leq \mathcal{C}^{*,\beta}.$$

Furthermore, Proposition 4.1.2 states that pure strategies $X \in \mathcal{A}_P(Q)$ give out the same cost functions for all three LOB models. That is to say

$$\inf_{X \in \mathcal{A}_P(Q)} \mathcal{C}^\infty(X) = \inf_{X \in \mathcal{A}_P(Q)} \mathcal{C}^\beta(X).$$

These three equations above imply that

$$\inf_{X \in \mathcal{A}_P(Q)} \mathcal{C}^\beta(X) \leq \inf_{X \in \mathcal{A}(Q)} \mathcal{C}^\beta(X).$$

At the same time, since $\mathcal{A}_P(Q) \subset \mathcal{A}(Q)$, we have

$$\inf_{X \in \mathcal{A}(Q)} \mathcal{C}^\beta(X) \le \inf_{X \in \mathcal{A}_P(Q)} \mathcal{C}^\beta(X).$$

Therefore, we have the equality $\inf_{X \in \mathcal{A}(Q)} \mathcal{C}^\beta(X) = \inf_{X \in \mathcal{A}_P(Q)} \mathcal{C}^\beta(X)$. $\square$

Proof of Proposition 5.2.3. Consider a round trip trading strategy: buy first $X^A = (0, \dots, 0, x_{t_i}, 0, \dots, 0)$, and then sell $X^B = (0, \dots, 0, x_{t_j}, 0, \dots, 0)$ with $t_i < t_j$ and $x_{t_i} = x_{t_j} > 0$. We can compute the same side and opposite side price impact dynamics and then obtain the cross-impact cost function

$$\begin{aligned}
\mathcal{C}^\beta(X^A, X^B) &= \frac{x_{t_i} x_{t_j}}{q(t_i)} e^{\rho(t_j - t_i)} \left(e^{-\beta(t_j - t_i)} - 1 \right) + \left(\frac{x_{t_i}^2}{2q(t_i)} + \frac{x_{t_j}^2}{2q(t_j)} \right) \\
&= \frac{m^2}{2q(t_i)} (b + 1 - 2a(1 - c)),
\end{aligned}$$

where we set $x_{t_i} = m > 0$ and $a = e^{-\rho(t_j - t_i)}$, $b = \frac{q(t_i)}{q(t_j)}$ and $c = e^{-\beta(t_j - t_i)}$.

Suppose that condition (5.2.1) does not hold i.e. $b < 2a(1-c) - 1$, the cross impact cost is then negative $\mathcal{C}^\beta(X^A, X^B) < 0$. Therefore, for any trading strategy $X \in \mathcal{A}(Q)$ there is

$$\inf_{X \in \mathcal{A}(0)} \mathcal{C}^\beta(X) \le \mathcal{C}^\beta(X^A, X^B) < 0.$$

This is a contradiction to the definition of absence of price manipulation strategy.

$\square$

Proof of Corollary 5.2.4. Adapting the notations from the Proof of Proposition 5.2.3, we have $a = e^{-\rho(t_j - t_i)}$, $b = \frac{q(t_i)}{q(t_j)}$ and $c = e^{-\beta(t_j - t_i)}$. Then condition (5.2.1) can be rewritten as $b \ge 2a - 1 - 2ac$. If $c \ge \frac{1}{2}$, and since $a < 1$, one obtains $b - 2a + 1 + 2ac \ge b + 1 - a > 0$. Therefore the assertion holds. $\square$

Proof of Proposition 5.3.1. Let us consider a buy program totally to buy $Q > 0$ shares. Consider a trading strategy: sell following $X^B = (0, \dots, 0, x_{t_i}, 0, \dots, 0)$, and then buy $X^A = (0, \ldots, 0, x_{t_j}, 0, ..., 0)$ with $t_i < t_j$. Here we set $x_{t_i} = m > 0$, $x_{t_j} = m + Q$, $a = e^{-\rho(t_j - t_i)}$, $b = \frac{q(t_i)}{q(t_j)}$, $c = e^{-\beta(t_j - t_i)}$ and

$$L(m) = m^2(b + 1 - 2a(1 - c)) + 2mQ(b - a(1 - c)) + bQ^2.$$

Then we compute the both sides price impact dynamics and then obtain the cross-impact cost function

$$\begin{aligned}
\mathcal{C}^\beta(X^A, X^B) &= \frac{x_{t_i} x_{t_j}}{q(t_i)} e^{-\rho(t_j - t_i)} \left(e^{-\beta(t_j - t_i)} - 1\right) + \left(\frac{x_{t_i}^2}{2q(t_i)} + \frac{x_{t_j}^2}{2q(t_j)}\right) \\
&= \frac{2m(Q + m)a(c - 1) + m^2}{2q(t_i)} + \frac{(Q + m)^2}{2q(t_j)} \\
&= \frac{1}{2q(t_i)} [m^2(b + 1 - 2a(1 - c)) + 2mQ(b - a(1 - c)) + bQ^2] \\
&= \frac{1}{2q(t_i)} L(m).
\end{aligned}$$

Look at the quadratic equation $L(m) = 0$. If the condition (5.3.2) does not hold in the sense that $\Delta := (b - a(1 - c))^2 - b(b + 1 - 2a(1 - c)) = a^2(1 - c)^2 - b > 0$, then the quadratic equation $L(m) = 0$ has two distinct solutions, which are given by

$$m_1 = Q\frac{\omega_1 - \sqrt{\Delta}}{m\omega_2}$$

and

$$m_2 = Q\frac{\omega_1 + \sqrt{\Delta}}{m\omega_2},$$

where $\omega_1 = a(1 - c) - b$ and $\omega_2 = b + 1 - 2a(1 - c)$.

- Case 1: if $\omega_2 > 0$.

Since $a < 1$ and $c < 1$, one has $0 < 1 - c < 1$ and then $a(1 - c) < 1$. The $\Delta > 0$ implies $b < a^2(1 - c)^2 < a(1 - c)$ which implies $\omega_1 > 0$. If $\omega_2 > 0$, we have $\omega_1 > \sqrt{\Delta}$ since

$$\omega_1^2 - \Delta^2 = (a(1 - c) - b)^2 - a^2(1 - c)^2 + b = b(b - 2a(1 - c) + 1) > 0.$$

Then the two solutions satisfies $0 < m_1 < m_2$. For any $m \in (m_1, m_2)$, by property of quadratic equation we have $\mathcal{C}^\beta(m) = \frac{L(m)}{2q(t_i)} < 0$. Thereafter, for any trading strategy $X \in \mathcal{A}(Q)$ there is

$$\inf_{X \in \mathcal{A}(Q)} \mathcal{C}^\beta(X) < \inf_{m \in (m_1, m_2)} \mathcal{C}^\beta(m) < 0.$$

In other words, by constructing a purchase trade strategy by selling quantity $m \in (m_1, m_2)$ first and then buying back $Q + m$, one could get negative liquidation cost.

- Case 2: $\omega_2 < 0$.

$\omega_2 < 0$ implies the necessary condition (5.2.1) for absence of PMS is violated. By the hierarchy Proposition 3.3.4, the cross-impact LOB model does not admit PLC. □

Proof of Proposition 5.3.2. The proof will be conducted in two steps. On one hand, let us prove that $q(s) \geq q(t)e^{-2\rho(t-s)}$ is the sufficient condition for the zero spread LOB model having positive liquidation costs.

The condition $q(s) \geq q(t)e^{-2\rho(t-s)}$ implies that for discrete times $\{t_0, t_1, ..., t_N\}$ one has $q(t_{i-1}) \geq q(t_i)e^{-\rho(t_i - t_{i-1})}$, or equivalently $a_i\tilde{a}_i < 1$ for $\forall i \in \{1, \ldots, N\}$, where $a_i = e^{-\rho(t_i - t_{i-1})}$ and $\tilde{a}_i = a_i \frac{q(t_i)}{q(t_{i-1})}$. Then, we can define vectors v_0 and v_i to be

$$v_0 = \frac{e_0}{\sqrt{q(t_0)}}$$

and

$$v_i = \tilde{a}_i v_{i-1} + \frac{e_i}{\sqrt{q(t_i)}} \sqrt{1 - a_i\tilde{a}_i},$$

where $1 \leq i \leq N$ and $e_0, \ldots, e_N$ are the canonical basis of $\mathbb{R}^{N+1}$. Note that v_i for $1 \leq i \leq N$ are linearly independent. We have the matrix $\Lambda_{ij} = \langle v_i, v_j \rangle$ as their Gram matrix. Thus Λ is positive definite and then the zero-spread cost function $\mathcal{C}^\infty \geq 0$.

On the other hand, let us prove that the zero spread LOB model having positive liquidation costs is the sufficient condition for a constant time varying shape function $q(t)$, satisfying $q(s) \geq q(t)e^{-2\rho(t-s)}$.

If the matrix Λ is positive definite, its minors

$$\det((\Lambda_{i,j})_{0\leq i,j\leq n}) = \frac{1}{q(t_0)} \prod_{i=1}^{n} \frac{1}{q(t_i)}(1 - a_i \tilde{a}_i), \quad 1 \leq n \leq N$$

are positive, which implies the condition (5.3.3).

□

Chapter 6

Conclusion

6.1 Concluding remarks

In this thesis, there were two main parts of research involved. First, in Chapter 2, the study was set up to explore the LOB resilience. The analysis was based on a market microstructure model proposed by Rosu [74]. Rather few investigations dealt with the resilience effect in the framework of a LOB from a game theoretical point of view. Second, in Chapter 3, 4 and 5, we worked on a new cross-impact LOB model by extending the LOB-based market impact model proposed by Obizhaeva and Wang [65] to include the two-side resilience. The difference between the existing LOB based market impact model and our cross-impact LOB model was the inclusion of the resilience effect on both bid and ask sides of the order book. In addition, our cross impact LOB model is the first to incorporate the bid-ask spread into the optimal execution strategy and market irregularity.

Chapter 2 verified order book resilience under a game theoretical model. We recapitulated the stochastic trading game in both the one-side case and the general case formulated in Rosu [74]. The main contribution of this chapter was that we proved the existence of the same side resilience in Proposition 2.2.3 in a mathematically rigourous way, by reinterpreting the same side resilience and the opposite side resilience in terms of the temporary price impact and permanent price impact. That was an essential step for verifying the existence of the resilience, since the price adjustment in this continuous time stochastic trading game is taking place instantaneously, while the resilience effect is a time-related feature of a LOB. We rewrote the solution for the recursive system of the sellers' utility function in Proposition 2.2.2 and then provided a proof of same side resilience in four steps by investigating the asymptotic behaviour of the difference of the price impact functions under the assumption of fast decaying arrival rates of impatient buyers who submitted more than one-unit order. We also constructed a counterexample of the same side resilient where the fast decaying assumption did not hold.

In Chapter 3, we formulate of the cross-impact LOB model. Apart from defining a trading strategy, a time-varying shape function and cost function, a crucial part of the model formulation was the introduction of the limit order regeneration on both ask and bid sides after a price shift created by market order. The same side price impact was described by processes D^A or D^B with a same side resilience rate ρ. The opposite side price impact was represented by processes L^A or L^B with a cross impact resilience rate β. The total price impacts were then linear combinations of the same side and opposite side price impacts. This new model successfully allowed an endogenous non-zero bid-ask spread in the LOB based market impact models. To the best of our knowledge, this model was the first one to model non-zero spread and two-side resilience effect in the market impact model framework. Moreover, in section 3.2 we indicated the generality of the cross-impact LOB model by replicating two existing LOB models under our model setting-ups. By sending the cross impact resilience rate β to infinity, we succeeded in replicating the zero-spread LOB model. The one-sided LOB model was also replicated here if the cross impact resilience rate β was set to be zero. At the end of this chapter, we provided three market irregularity definitions. We proved that for cross-impact LOB models, the absence of transaction-triggered price manipulation (TTPM) implied the positivity of the liquidity costs (PLC) and the positivity of liquidity cost lead to the absence of price manipulation strategy (PMS) in Proposition 3.3.4.

In Chapter 4 we applied the cross-impact LOB model for studying the optimal execution problem. We proved the existence of optimal execution strategy under the cross-impact LOB framework in four cases: model of volume impact reversion in discrete time, model of volume impact reversion in continuous time, model of price impact reversion in discrete time and model of price impact reversion in continuous time. We found that for any strategy in the admissible set, the zero-spread cost is a lower bound of the cross-impact cost. Furthermore, we proved that the minimum costs of the zero-spread LOB model is a lower bound of the minimum cost of the cross-impact LOB model. The relationship between cost functions of zero-spread LOB model and cross-impact LOB model was presented in Proposition 4.1.2 and Corollary 4.1.3. With the help of the costs relationship between models, we transferred the problem of proving existence of optimal solution under the cross-impact LOB model to the existence problem under the zero-spread LOB model. For discrete time cases, we proved the coerciveness of cost functions and solved the optimal strategy by the method of Lagrange multiplier. Then we took the continuous analog of the discrete time optimal strategy and proved the optimality via verification argument. A by-product of this process was that we obtained sufficient conditions for absence of PMS and TTPM under the zero-spread LOB model. They were summarised in Table 4.1. Our results contributed to the zero-spread LOB model literature by generalising the optimal trading strategy and absence conditions of market irregularities since a general form of shape function was used. Under the assumption of a constant time-varying shape function, we presented some numerical examples of optimal strategies under the cross-impact LOB model. As a result, the relationships between respectively the optimal strategy and the depth function, the cross impact resilience rate, the same side resilience rate were explicitly investigated. Figure 4.4 suggested that as the cross impact resilience rate increased, more opposite side orders were used under the LOB model with TTPM. Figure 4.6 suggested that as the same side resilience rate increased, the optimal strategies were less volatile.

In Chapter 5 we addressed the regularity problems of market impact model when one cannot obtain closed-form optimal solutions of the execution problem like in Chapter 4. We answered the questions raised by Fruth et al. [31]: might it be true that the LOB model with non-zero spread exclude these price manipulations. We found that the cross-impact LOB models possess non-zero spread, but it is still profitable to manipulate prices. We proved that the profitability of the price manipulation strategies depends on the resiliency of the order book and the dynamics of the shape function. The conditions of the absence of three market irregularities were presented in Proposition 5.1.1, Proposition 5.1.2, Proposition 5.2.1, Proposition 5.3.1 and Proposition 5.3.2. According to these propositions, we concluded that market irregularity conditions under zero-spread LOB models were weaker than those under the cross-impact LOB models.

6.2 Further research

We have argued in Chapter 1 and Chapter 2 that there are both the same side resilience and opposite side resilience in the limit order book market, nevertheless, there are much less theoretical studies on the resilience effect than on other features of the order book. The game theoretical models are still introducing a rather large amount of free parameters, most of which cannot be measured directly. We may resort to the stochastic models of the limit order book. It would be desirable to be able to model the dynamic bid-ask spread endogenously, via the various order flows of the trading within. More market microstructure research are needed to help with explaining and identifying the after-shock price formation.

From perspective of optimal execution problem, it is certainly a challenge to come up with a mathematically tractable model that cover all three layers of order splitting. A first step could be to add the choice of limit order into the optimal execution strategy. Apart from the studies about the second layer execution discussed in Chapter 1, we believe that the combination usage of limit orders and/or dark liquidity inside the spread can be a research direction. Indeed, the resilience effect is the process of regeneration of limit orders inside the after-price-shock spread.

There is still a gap for this optimisation problem.

Bibliography

[1] F. Abergel and A. Jedidi. A mathematical approach to order book modelling. In F. Abergel, B.K. Chakrabarti, A. Chakraborti, and M. Mitra, editors, *Econophysics of Order-driven Markets*, New Economic Windows, pages 93–107. Springer Milan, 2011.

[2] A. Alfonsi and J. I. Acevdeo. Optimal execution and price manipulations in time-varying limit order books. *Applied Mathematical Finance*, Early View, 2013.

[3] A. Alfonsi and A. Schied. Optimal execution and absence of price manipulations in limit order book models. *SIAM Journal on Financial Mathematics*, 1:490–522, 2010.

[4] A. Alfonsi, A. Fruth, and A. Schied. Constrained portfolio liquidation in a limit order book model. *Banach Centre Publications*, 83:9–25, 2008.

[5] A. Alfonsi, A. Fruth, and A. Schied. Optimal execution strategies in limit order books with general shape functions. *Quantitative Finance*, 10(2):143–157, 2010.

[6] A. Alfonsi, A. Schied, and A. Slynko. Order book resilience, price manipulation, and the positive portfolio problem. *SIAM Journal on Financial Mathematics*, 3(1):511–533, 2012.

[7] F. Allen and G. Gorton. Stock price manipulation, market microstructure and asymmetric information. *European Economic Review*, 36:624–630, 1992.

[8] R. Almgren. Optimal execution with nonlinear impact functions and trading-enhanced risk. *Applied Mathematical Finance*, 10(1):1–18, 2003.

[9] R. Almgren and N. Chriss. Optimal execution of portfolio transactions. *Journal of Risk*, 3(2):5–39, 2000.

[10] M. Avellaneda and S. Stoikov. High-frequency trading in a limit order book. *Quantitative Finance*, 8(3):217–224, 2008.

[11] E. Bayraktar and M. Ludkovski. Liquidation in limit order books with controlled intensity. *Mathematical Finance*, Early View, 2012.

[12] J. Bergin and W. B. Macleod. Continuous time repeated games. *International Economic Review*, 34(1):21–37, 1993.

[13] D. Bertsimas and A. W. Lo. Optimal control of execution costs. *Journal of Financial Markets*, 1(1):1–50, 1998.

[14] B. Biais, P. Hillion, and C. Spatt. An empirical analysis of the limit order book and the order flow in the Paris Bourse. *The Journal of Finance*, 50(5): 1655–1689, 1995.

[15] B. Biais, L. Glosten, and C. Spatt. Market microstructure: A survey of microfoundations, empirical results and policy implications. *Journal of Financial Markets*, 8(2):217–264, 2005.

[16] R. Bloomfield, M. O'Hara, and G. Saar. The "make or take" decision in an electronic market: Evidence on the evolution of liquidity. *Journal of Financial Economics*, 75(1):165–199, 2005.

[17] J. P. Bouchaud, Y. Gefen, M. Potters, and M. Wyart. Fluctuations and response in financial markets: the subtle nature of 'random' price changes. *Quantitative Finance*, 4:176–190, 2004.

[18] J. P. Bouchaud, J. D. Farmer, and F. Lillo. How markets slowly digest changes in supply and demand. In T. Hens and K. R. Schenk-Hoppe, editors, *Handbook*

of Financial Markets: Dynamics and Evolution, pages 57–160. North-Holland, 2009.

[19] M. K. Brunnermeier and L. H. Pedersen. Predatory trading. *The Journal of Finance*, 60(4):1825–1864, 2005.

[20] S. Buti and B. Rindi. Undisclosed orders and optimal submission strategies in a limit order market. *Journal of Financial Economics*, 109(3):797–812, 2013.

[21] A. Cartea, S. Jaimungal, and J. Ricci. Buy low sell high: a high frequency trading perspective. *SIAM Journal on Financial Mathematics*, Forthcoming.

[22] G. Cebiroğlu and U. Horst. Optimal display of iceberg orders. Working paper, Humboldt University of Berlin, 2011.

[23] W. G. Christie and P. H. Schults. Why do NASDAQ market makers avoid odd-eight quotes? *The Journal of Finance*, 49:1813–1840, 1994.

[24] R. Cont and A. Kukanov. Optimal order placement in limit order markets. Working paper series, Imperial College London, Columbia University, 2012.

[25] R. Cont and A. de Larrard. Price dynamics in a markovian limit order market. *SIAM Journal on Financial Mathematics*, 4(1):1–25, 2013.

[26] F. Cucker and A. G. Corbalan. An alternate proof of the continuity of the roots of a polynomial. *The American Mathematical Monthly*, 96(4):342–345, 1989.

[27] H.A. Degryse, F.de Jong, M.van Ravenswaaij, and G. Wuyts. Aggressive orders and the resiliency of a limit order market. *Review of Finance*, 9(2):201–242, 2005.

[28] D. Easley and M. O'Hara. Price, trade size, and information in securities markets. *Journal of Financial Economics*, 19(1):69–90, 1987.

[29] A. Esser and B. Mönch. The navigation of an iceberg: The optimal use of hidden orders. *Finance Research Letters*, 4(2):68–81, 2007.

[30] T. Foucault, O. Kadan, and E. Kandel. Limit order book as a market for liquidity. *The Review of Financial Studies*, 18(4):1171–1217, 2005.

[31] A. Fruth, T. Schöneborn, and M. Urusov. Optimal trade execution and price manipulation in order books with time-varying liquidity. *Mathematical Finance*, Early View, 2013.

[32] J. Gatheral. No-dynamic-arbitrage and market impact. *Quantitative Finance*, 10(7):749–759, 2010.

[33] J. Gatheral. Optimal order execution. In *JOIM Fall Conference, Boston*, 2011.

[34] J. Gatheral and A. Schied. Optimal trade execution under geometric Brownian motion in the Almgren and Chriss framework. *International Journal of Theoretical and Applied Finance*, 14(3):353–368, 2011.

[35] J. Gatheral and A. Schied. Dyanmical models of market impact and algorithms for optimal order execution. In Jean-Pierre Fouque and J. A. Langsam, editors, *Handbook on Systemic Risk*, pages 579–602. Cambridge University Press, 2013.

[36] J. Gatheral, A. Schied, and A. Slynko. Exponential resilience and decay of market impact. In F. Abergel, B. K. Chakrabarti, A. Chakraborti, and M. Mitra, editors, *Econophysics of Order-driven Markets*, New Economic Windows, pages 225–236. Springer, 2011.

[37] J. Gatheral, A. Schied, and A. Slynko. Transient linear price impact and fredholm integral equations. *Mathematical Finance*, 22(3):445–474, 2012.

[38] L. R. Glosten and P. R. Milgrom. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14(1):71–100, 1985.

[39] C. A. Goettler, R. L.and Parlour and U. Rajan. Equilibrium in a dynamic limit order market. *The Journal of Finance*, 60(5):2149–2192, 2005.

[40] M. D. Gould, M. A. Porter, S. Williams, M. McDonald, D. J. Fenn, and S. D. Howison. Limit order books. *Quantitative Finance*, 13(11):1709–1742, 2013.

[41] O. Guéant and C.-A. Lehalle. General intensity shapes in optimal liquidation. *Mathematical Finance*, Early view, 2013.

[42] O. Guéant, C.-A. Lehalle, and J. Fernandez-Tapia. Optimal portfolio liquidation with limit orders. *SIAM Journal on Financial Mathematics*, 3(1):740–764, 2012.

[43] F. Guibaud and H. Pham. Optimal high-frequency trading in a pro rata microstructure with predictive information. *Mathematical Finance*, Early view, 2013.

[44] F. Guilbaud and H. Pham. Optimal high frequency trading with limit and market orders. *Quantitative Finance*, 13(1):79–94, 2013.

[45] L. Harris. Optimal dynamic order submission strategies in some stylized trading problems. *Financial Markets, Institutions and Instruments*, 7(2):1–76, 1998.

[46] J. B. Hiriart-Urruty and A. Seeger. A variational approach to copositive matrices. *SIAM Review*, 52(4):593–629, 2010.

[47] B. Hollifield, R. A. Miller, and P. Sandas. Empirical analysis of limit order markets. *The Review of Economic Studies*, 71(4):1027–1063, 2004.

[48] G. Huberman and W. Stanzl. Price manipulation and quasi-arbitrage. *Economitrica*, 72(4):1247–1275, 2004.

[49] G. Huberman and W. Stanzl. Optimal liquidity trading. *Review of Finance*, 9 (2):165–200, 2005.

[50] R. Huitema. Optimal portfolio execution using market and limit orders. Working paper series, University of Zrich, 2014.

[51] I. Kharroubi and H. Pham. Optimal portfolio liquidation with execution cost and risk. *SIAM Journal on Financial Mathematics*, 1(1):897–931, 2010.

[52] F. Klöck. Regularity of market impact models with stochastic price impact. Working paper series, University of Mannheim, 2012.

[53] F. Klöck, A. Schied, and Y. Sun. Existence and absence of price manipulation in a market impact model with dark pool. Preprint, University of Mannheim, Cornell University, ORIE, 2014.

[54] P. Kovaleva and G. Iori. Optimal trading strategies in a limit order market with imperfect liquidity. Working paper series, City University, London, 2012.

[55] P. Kratz and T. Schöneborn. Optimal liquidation in dark pool. EFA 2009 bergen meetings paper, Humboldt Universitat of Berlin and Deutsche Bank AG, London, 2013.

[56] A. S. Kyle. Continuous auctions and insider trading. *Econometrica*, 53(6): 1315–1335, 1985.

[57] S. Laruelle, C.-A. Lehalle, and G. Pagès. Optimal split of orders across liquidity pools: A stochastic algorithm approach. *SIAM Journal on Financial Mathematics*, 2(1):1042–1076, 2011.

[58] C.-A. Lehalle. Market microstructure knowledge needed for controlling an intra-day trading process. In J.-P. Fouque and J. A. Langsam, editors, *Handbook on Systemic Risk*, pages 549–578. Cambridge University Press, 2013.

[59] C.-A. Lehalle, Q. Li, and U. Horst. Optimal trading in a two-sided limit order book. Working paper series, University of Évry-Val d'Essonne, 2013.

[60] C. Lorenz and A. Schied. Drift dependence of optimal order execution strategies under transient price impact. *Finance and Stochastics*, 17(4):743–770, 2013.

[61] J. Lorenz and R. Almgren. Mean-variance optimal adaptive execution. *Applied Mathematical Finance*, 18(5):395–422, 2011.

[62] A. Madhavan. Market microstructure: A survey. *Journal of Financial Markets*, 3(3):205–258, 2000.

[63] E. Maskin and J. Tirole. Markov perfect equilibrium, I: observable actions. *Journal of Economic Theory*, 100(2):191–219, 2001.

[64] F. Naujokat and N. Westray. Curve following in illiquid markets. *Mathematics and Financial Economics*, 4(4):299–335, 2011.

[65] A. Obizhaeva and J. Wang. Optimal trading strategy and supply/demand dynamics. *Journal of Financial Markets*, 16(1):1–32, 2013.

[66] M. O'Hara. *Market Microstructure Theory*. Wiley-Blackwell Publishing, 1997.

[67] C. A. Parlour. Price dynamics in limit order markets. *The Review of Financial Studies*, 11(4):789–816, 1998.

[68] C. A. Parlour and D. J. Seppi. Limit order market: a survey. In A. Thakor and A. Boot, editors, *Handbook of financial intermediation and banking*, pages 63–96. Elsevier, 2008.

[69] A. Ponzi, F. Lillo, and R. N. Mantegna. Market reaction to temporary liquidity crisis and the permanent market impact. Working paper series, INFM-CNR, University of Palermo, Santa Fe Institute, 2008.

[70] M. Potters and J. P. Bouchaud. More statistical properties of order books and price impact. *Physica A: Statistical Mechanics and its Applications*, 324(1): 133–140, 2003.

[71] S. Predoiu, G. Shaikhet, and S. Shreve. Optimal execution in a general one-sided limit order book. *SIAM Journal on Financial Mathematics*, 2(1):183–212, 2011.

[72] R. Roll. A simple implicit measure of the effective bid-ask spread in an efficient market. *The Journal of Finance*, 39(4):1127–1139, 1984.

[73] I. Rosu. Multi-stage game theory in continuous time. Working paper series, University of Chicago, 2006.

[74] I. Rosu. A dynamic model of the limit order book. *The Review of Financial Studies*, 22(11):4601–4641, 2009.

[75] I. Rosu. Liquidity and information in limit order markets. Working paper series, HEC Paris, 2014.

[76] A. Schied. Robust strategies for optimal order execution in the Almgren-Chriss framework. *Applied Mathematical Finance*, 20(3):264–286, 2013.

[77] A. Schied and T. Schöneborn. Risk aversion and the dynamics of optimal liquidation strategies in illiquid markets. *Finance and Stochastic*, 13(2):181–204, 2009.

[78] A. Schied and A. Slynko. Some mathematical aspects of market impact modelling. In J. Blath, P. Imkeller, and S. Roelly, editors, *Surveys in Stochastic Processes*, pages 153–179. EMS Series of Congress Reports, 2011.

[79] A. Schied, T. Schöneborn, and M. Tehranchi. Optimal basket liquidation for CARA investors is deterministic. *Applied Mathematical Finance*, 17(6):471–489, 2010.

[80] E. Smith, J. D. Farmer, L. Gillemot, and S. Krishnamurthy. Statistical theory of the continuous double auction. *Quantitative Finance*, 3(6):481–514, 2003.

[81] H. R. Stoll. The pricing of security dealer services: an empirical study of NASDAQ stocks. *The Journal of Finance*, 33(4):1153–1172, 1978.

[82] I. M. Toke. "market making" in an order book model and its impact on the spread. In F. Abergel, B. K. Chakrabarti, A. Chakraborti, and M. Mitra, editors, *Econophysics of Order-driven Markets*, New Economic Windows, pages 49–64. Springer, 2011.

[83] B. Tóth, J. Kertész, and J. D. Farmer. Studies of the limit order book around large price changes. *The European Physical Journal B*, 71(4):499–510, 2009.

[84] H. Väliaho. Criteria for copositive matrices. *Linear algebra and its applications*, 81:19–34, 1986.

[85] P. Weber and B. Rosenow. Order book approach to price impact. *Quantitative Finance*, 5(4):357–364, 2005.

[86] A. Weiss. Executing large orders in a microscopic market model. Preprints, DFG Research Centre MATHEON, 2010.

[87] M. Wyart, J. P. Bouchaud, J. Kockelkoren, M. Potters, and M. Vettorazzo. Relation between bid-ask spread, impact and volatility in order-driven markets. *Quantitative Finance*, 8(1):41–57, 2008.